Technical Report 1056

Terrain Appreciation in Virtual Environments: Spatial Knowledge Acquisition

Michael J. Singer
U.S. Army Research Institute

Robert C. Allen
Daniel P. McDonald
John P. Gildea
University of Central Florida, Orlando

February 1997

United States Army Research Institute for the Behavioral and Social Sciences

U.S. ARMY RESEARCH INSTITUTE FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency Under the Jurisdiction of the Deputy Chief of Staff for Personnel

EDGAR M. JOHNSON
Director

Technical review by

Gene W. Fober
Larry L. Meliza

NOTICES

DISTRIBUTION: Primary distribution of this report has been made by ARI. Please address correspondence concerning distribution of reports to : U.S. Army Research Institute for the Behavioral and Social Sciences, ATTN: PERI-STP, 5001 Eisenhower Ave., Alexandria, Virginia 22333-5600

FINAL DISPOSITION: This report may be destroyed when it is no longer needed. Please do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: The findings in this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

REPORT DOCUMENTATION PAGE

1. REPORT DATE (dd-mm-yy) 15-01-97	2. REPORT TYPE Technical Report	3. DATES COVERED (from. . . to) 7/94 to 1/97
4. TITLE AND SUBTITLE Terrain Appreciation in Virtual Environments: Spatial Knowledge Acquisition		5a. CONTRACT OR GRANT NUMBER
		5b. PROGRAM ELEMENT NUMBER PE62785A
6. AUTHOR(S) Singer, Michael J. (Army Research Institute) Allen, Robert, C., McDonald, Daniel P., & Gildea, John, P. (University of Central Florida; Orlando FL)		5c. PROJECT NUMBER PRA791
		5d. TASK NUMBER TA 2111
		5e. WORK UNIT NUMBER H01
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) U.S. Army Research Institute for the Behavioral and Social Sciences ATTN: PERI- 5001 Eisenhower Avenue Alexandria, VA 22333-5600		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Research Institute for the Behavioral and Social Sciences 5001 Eisenhower Avenue Alexandria, VA 22333-5600		10. MONITOR ACRONYM ARI
		11. MONITOR REPORT NUMBER

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT *(Maximum 200 words)*: The U. S. Army Research Institute is investigating the requirements for using Virtual Environments (VE) in training dismounted soldiers. This experiment investigated the effects of different VE parameters on spatial knowledge acquisition by comparing learning in advanced VE, restricted VE, and standard map training. This report also provides information about VE displays, head-coupling, presence, and simulator sickness associated with spatial knowledge acquisition in VE. The activities used during the learning phase of the experiment are generic to dismounted soldier activities.

The high level virtual environment (Hi-VE) condition had a Stereoscopic Head Mounted Display (HMD) with fully head-coupled gaze control, and treadmill-based movement control. The restricted VE configuration (Low-VE) used the same HMD with both gaze direction and viewpoint movement controlled by a joystick. The map training participants used expanded topographical maps and were subsequently tested in the Hi-VE configuration. Participants were all trained on the definitions and representational configuration of a reduced set of topographical features, and dismissed if unable to reach a minimum criterion. The Simulator Sickness Questionnaires (SSQ), and the Immersive Tendencies Questionnaire (ITQ) were administered before the VE experience. Participants received training in VE movement and control before the experimental

15. SUBJECT TERMS: Virtual Environments; Simulation Fidelity; Spatial Knowledge Acquisition;Simulator Sickness; Presence.

SECURITY CLASSIFICATION OF					
16. REPORT Unclassified	17. ABSTRACT Unclassified	18. THIS PAGE Unclassified	19. LIMITATION OF ABSTRACT Unclassified	20. NUMBER OF PAGES	21. RESPONSIBLE PERSON (Name and Telephone Number) Dr. Michael J. Singer (407) 384-3993

14. ABSTRACT. (continued)

training and testing was conducted. The SSQ was repeated, and a Presence Questionnaire (PQ) was administered after the experimental session.

The results demonstrate that better spatial knowledge is gained through more highly-interactive Virtual Environments experiences than through equivalent practice with topographical maps. The more normal interactions and cues supported by the virtual environment seems to support better spatial recognition and knowledge of specific landmarks than can be acquired through purely cognitive exercises with symbolic (topographical) representations of terrain. The longer times in the VE, and the different levels of VE equipment did not lead to differences in simulator sickness. The amount of time spent in the VE was significantly positively correlated with several measures of landmark direction accuracy.

Technical Report 1056

Terrain Appreciation in Virtual Environments: Spatial Knowledge Acquisition

Michael J. Singer
U.S. Army Research Institute

Robert C. Allen, Daniel P. McDonald, and John P. Gildea
University of Central Florida, Orlando

Simulator Systems Research Unit
Stephen L. Goldberg, Chief

U.S. Army Research Institute for the Behavioral and Social Sciences
5001 Eisenhower Avenue, Alexandria, Virginia 22333-5600

Office, Deputy Chief of Staff for Personnel
Department of the Army

February 1997

Army Project Number
20262785A791

Education and Training Technology

FOREWORD

The U.S. Army has made a substantial commitment to Distributed Interactive Simulation (DIS) and the electronic battlefield for training, readiness, concept development, and test and evaluation. The current DIS training system, Simulation Network, and the next generation system, the Close Combat Tactical Trainer, are both designed to provide realistic training for platform-based warfighting. These systems are not designed to provide training for individual dismounted soldiers. Virtual Environment (VE) technology completely replaces one or more real-world sensory domains with computer-generated simulation. VE technology has the potential to provide Individual Combatant Simulations for interaction with the electronic battlefield, as well as for individual training. One area of challenging research is identifying and quantifying the effects of VE system characteristics on learning, skill acquisition, retention, and transfer of U.S. Army tasks.

This report describes an experiment in an ongoing series addressing VE technology for training dismounted soldiers. The experiment described here was designed to address the effect of VE interactivity level on spatial knowledge acquisition in VEs. Spatial knowledge learning was selected as a research focus because it is basic to many different soldier activities that will eventually be incorporated in VE-based training. The results of this experiment indicate that better spatial knowledge is gained through more highly interactive VE experiences than through equivalent practice with topographical maps. The findings from this research can be used to select the VE characteristics needed for effective spatial knowledge acquisition and other spatially based tasks learned or practiced in VEs.

The U.S. Army Research Institute for the Behavioral and Social Sciences' Simulator Systems Research Unit conducts research with the goal of providing information that will improve the effectiveness of training simulators and simulations. The work described here is a part of ARI Research Task 2111, VIRTUE - Virtual Environments for Combat Training and Mission Rehearsal.

ZITA M. SIMUTIS
Technical Director

EDGAR M. JOHNSON
Director

ACKNOWLEDGMENT

This research could not have been conducted without the programming and technical support provided by the Institute for Simulation and Training, University of Central Florida. We appreciate the efforts of Kimberly Parsons, Kevin Mueller, Jim Parsons, Sunil Dixit, and Greg Wiatroski in producing the experimental virtual environment and the data collection software and hardware.

TERRAIN APPRECIATION IN VIRTUAL ENVIRONMENTS: SPATIAL KNOWLEDGE ACQUISITION

EXECUTIVE SUMMARY

Research Requirement:

The U.S. Army has committed to using Distributed Interactive Simulation (DIS) and the electronic battlefield for training, concept development, testing, and evaluation. Current and developing systems are designed to provide training for soldiers fighting from vehicles, but are not designed to provide realistic training for dismounted infantry. Virtual Environment (VE) technology presents a new way to simulate real world activities, which will allow the U.S. Army to conduct planning, training, and rehearsal activities for both individual and collective dismounted soldier tasks. Basic to these efforts is the common context for individual combatants who need to move, observe, shoot, and communicate. A fundamental spatial knowledge of the terrain and position is required as a basis for these and other soldier tasks and activities. Research in spatial knowledge acquisition is the first step in establishing the benefits and deficiencies of training and rehearsing complex soldier activities and tasks in VE.

Procedure:

In this experiment, three groups of participants "moved" through simulated terrain, performing simple cognitive terrain appreciation activities. The High-Level VE (Hi-VE) group walked on an instrumented treadmill while wearing a helmet-mounted display and using a pointing wand for indicating directions or locations and selecting objects. The Low-level VE (Lo-VE) group moved through the same simulated terrain, performing the same activities. This group was seated while observing the terrain through the same helmet-mounted display without head-tracking (with a fixed view), using a joystick to control view/movement, and the same pointing wand for direction indication or selections. The control group (Map) performed the activities using topographical maps, with paced study replacing movement through terrain, while seated at a desk. After the practice session, during which participants in all conditions followed the same designated route and performed the same terrain appreciation tasks, participants' configuration knowledge of the terrain was tested. The VE conditions were tested in the same condition in which they practiced, and the Map condition transferred to the Hi-VE configuration.

Findings:

The participants trained with more highly interactive VE experiences developed significantly better spatial knowledge than participants trained in comparable map exercises. This difference holds over different terrains, although there was significant variation over different test sites.

Measures of projective convergence, which combine direction and distance estimations to produce average accuracy and consistency measures, found a significant difference between the VE conditions and the Map practice in the consistency of landmark identification in one terrain. Significant correlations between correct direction indication of visible and non-visible landmarks indicates that positionally related spatial knowledge develops quite early in acquisition. A significant negative correlation was also found between participants' reports of their amount of previous experience with VE and the change in reported simulator sickness symptoms. This may indicate that increasing experience with different VE configurations might lead to decreasing problems with simulator sickness.

Utilization of Findings:

The U.S. Army will employ VE technology for training, mission rehearsal, planning, and new equipment concept testing. Understanding the improved acquisition of spatial knowledge from VE experiences will enable specification of VE configurations for different kinds of training and mission rehearsal. The results reported here begin to provide trainers and leaders with a basis for planning training and mission exercises before deployment. In addition, the finding that greater VE experience leads to less change in simulator sickness symptom levels indicates that soldiers with extensive VE training might not suffer from simulator sickness problems. This issue is one that should be pursued through ancillary questionnaires in subsequent VE research and validation.

TERRAIN APPRECIATION IN VIRTUAL ENVIRONMENTS: SPATIAL KNOWLEDGE ACQUISITION

CONTENTS

CONTENTS (Continued)

Page

LIST OF TABLES

TERRAIN APPRECIATION IN VIRTUAL ENVIRONMENTS: SPATIAL KNOWLEDGE ACQUISITION

Perhaps the most critical problem in training has always been how to represent operational reality within the training situation (Hays & Singer, 1989). Typically this representation has been done through some form of simulation, and many different configurations can be used as training devices. Virtual Environment (VE) technology is the label applied to a new generation of computer-driven simulations of both the real and synthetic worlds. Virtual Environment (VE) technology is used here to mean an equipment configuration that completely replaces the input from at least one sensory domain (typically vision). The prototypical VE configuration uses helmet mounted displays to replace the normal visual field with computer generated visual information, and may be combined with joysticks, gloves, or other apparatus to control movement through or interact with that environment. The key is that, as with reality, one can interact with objects in the simulated environment from user-chosen points of view. The interaction between the user/trainee and the environment can be used as a learning trial, a practice trial, verification of a simulated scenario, or even validation of some new equipment configuration

It seems obvious that VE technology presents new ways to simulate real world soldier activities, which will allow the U.S. Army to expand Distributed Interactive Simulation (DIS) to include more than simple, non-interactive representations of dismounted infantry. The U.S. Army Research Institute for the Behavioral and Social Sciences (ARI), Simulator Systems Research Unit (SSRU) at Orlando, Florida, has an ongoing research program on the use of VEs in training, called VIRTUE. An initial study was done in 1992 on the state of VE technology for soldier interaction as a basis for planning this research program (Jacobs, Crooks, Crooks, Colburn, Fraser, Gorman, Madden, Furness, and Tice, 1994). The study investigated the type and level of VE requirements for replicating the entire range of normal soldier interactions with the real world. That study concluded that available technology was not currently sufficient for exactly replicating the real world sufficiently for normal soldier tasks. This conclusion raises the decades old simulation question of "How much is enough?" In other words, how much can a soldier learn from a less than perfect simulation that is relevant to soldiers' real world activity (Hays and Singer, 1989).

The overall focus of the SSRU VIRTUE program is the investigation of VE technology for learning and effective transfer of dismounted soldier activities, and the inclusion of individual soldier activities in DIS combat exercises (Witmer, Bailey, Knerr, & Abel, 1994). These activities consist of planning, training, and rehearsal activities for both individual and collective soldier tasks. The primary interest is in small-group leader tasks, subtasks, and activities (for example platoon, squad, or fire team leaders). Especially important are the evaluative and decision-making skills that provide the fundamental skills for an effective warfighter. Basic to all of these efforts is the common context of individual combatants who need to move, observe, shoot, and communicate. The research program has been investigating whether these four basics

can be adequately practiced and skills improved through the use of VE technology. A common belief in the VE literature is that the context, workload, and cues for more complex skills can be provided in current or emerging VE systems. Providing more realism for training and practicing the basic activities and more complex skills (for example, decision-making) will presumably engender transfer, and transfer of these skills is crucial to mission performance.

Terrain Appreciation

A review of dismounted Infantry Army Training and Evaluation Program (ARTEP) elements was conducted early in the VIRTUE program (Jacobs, Crooks, Crooks, Colburn, Fraser, Gorman, Madden, Furness, & Tice, 1994). The review identified major activities that could be performed, trained, or practiced in VE (Jacobs, et al., 1994). They analyzed the critical stimuli required in the performance of the major activities, then evaluated the practicality of using VE technology to provide critical stimuli. There were three areas of primary consideration in that analysis. One area focused on the sensory modalities used in the performance of the activities. Another analysis addressed the projected capability of VE technology to present the necessary stimuli (circa 1992, when these analyses were performed). The third area assessed the potential for performance transfer from VE to the real world.

Land Navigation

There were many terrain interaction activities that had high combined rankings in terms of good estimated cost effectiveness, projected transfer effectiveness, current technological capability, and commonality across a large number of ARTEPS. These activities include Identifying Safe and Danger Areas, Movement by Direction, and Identifying Overwatch Positions (Jacobs, et al., 1994). Underlying all of these activities is the interaction of terrain appreciation skills and spatial knowledge of the operational terrain possessed by the soldier. Terrain appreciation means having a general understanding of how to use terrain features in performing soldier tasks such as weapons emplacement, defensive positions, and land navigation. The application of terrain appreciation requires knowing the configuration of terrain through experience and/or from map study, in order to use that spatial representation to plan and perform military activities.

The primary objective of this experiment was to begin investigating the effectiveness of VE configurations in the acquisition and maintenance of spatial knowledge in the context of terrain appreciation. There are several tasks that fall into the category of activities called terrain appreciation and all of these tasks are based in an understanding of the terrain. Basic terrain appreciation involves learning to apply general tactical rules about terrain features in the performance of soldier tasks (e.g., determining threat vectors to troop positions, determining optimal weapons emplacement for attack or defense). Another major aspect of terrain appreciation is the ongoing acquisition of spatial knowledge, an awareness of one's location and the configuration of the surrounding terrain. This is particularly important for soldiers on the

battlefield, as location provides a basic referent for many important soldier tasks. The platoon leader, NCO, and Forward Observer must all know where they are in order to be able to issue correct commands for indirect fire. In addition, leaders must know where they are, where they need to be, and how to get there, in order to be able to redirect movement in response to changed tactical orders (SSGT. Shonkwiler, Mr. McIlroy, personal communications, 1995).

Spatial Knowledge

As all organisms demonstrate learning and memory for environmental arrangements, the existence of some organized form of spatial knowledge has long been presumed. Working from that assumption, researchers have begun to investigate the representational structure of spatial knowledge and conditions of acquisition (Goldin & Thorndyke, 1981; Siegal & White, 1975). There have been differing arguments advanced for the representation format that humans use in storing and manipulating learned information (Eysenck & Keane, 1990). It is almost certain that spatial knowledge structures are encoded in a fashion that allows the coordination of semantics and imagery in dealing with spatial aspects of environmental information (Paivio, 1986). Spatial memory probably is also built upon increasing levels of elaboration (Craik & Lockhart, 1972; Craik & Tulving, 1976). Goldin and Thorndyke (1981) described three general levels of spatial knowledge as a result of their spatial memory research. In thier general schema, the base level is knowledge about individual landmarks, what the landmark characteristics are and the orientation of one landmark to another. The intermediate level is comprised of knowledge about routes between landmarks. The highest level of organization is survey or configuration which relates landmarks and routes into a connected map that allows the generation of new routes with some level of accuracy (Goldin & Thorndyke, 1981; Siegal & White, 1975; Witmer, Bailey, & Knerr, 1995).

It makes sense that the lowest level would be knowledge about landmarks (and is labeled landmark by Goldin & Thorndyke, 1981; 1982). Landmark knowledge includes the outstanding visual characteristics, the feature's orientation to common directions (e.g., East or West), and environmental context. Distinctiveness of the perceptual features of the landmark probably forms the primary set or layer of cues. The context of the landmark can also contribute to the distinctiveness of the feature. For example, a rounded hill in a panorama of same-sized rounded hills is not very distinctive, and hence is less useful as a landmark for organizing spatial memory. A hill that is higher, has a sharper slope, distinctive coloration, or atypical rock formations at the crest would provide more salience as a landmark. Finally, some semantic and organizational information might be associated with the imagery-based representation of the landmark, such as that it is forty degrees East of some other landmark, or North of an orienting feature like a road. Route knowledge (called procedural knowledge by Goldin & Thorndyke, 1981; 1982) is the organized sequence of landmarks, orientations between landmarks, and distances between landmarks and locations that provide or enable a transition from here to there. Some research has shown that active interaction with the environment enhances knowledge of routes over that acquired from passive movement through the environment or study of maps (Goldin &

Thorndyke, 1981; 1982). This same research shows that the acquisition of coherent spatial knowledge is not all or none, but incremental. This means that *some* knowledge of landmark sequences can just as easily be acquired by film or map study as by experience, although interaction with the environment seems to enhance the orientation aspects of route knowledge. Configuration knowledge (using the label introduced by Witmer, Bailey, & Knerr, 1995; referred to as survey knowledge by Goldin & Thorndyke, 1981) is required to be able to relate distances and orientations between locations, landmarks, or routes accurately, and especially to create new routes. As noted above, experience-based learning has been shown to be better for acquiring configuration knowledge (tested by reporting distances and orientation to unseen locations) than map learning (Goldin & Thorndyke, 1981; 1982).

Spatial Knowledge Acquisition in Virtual Environments

Previous research has demonstrated that some level of spatial knowledge can be acquired in VE and transfered to the real world. Regian, Shebliske, and Monk (1992) examined cognitive learning and representation of a "large scale space" from a VE, and tested in the same VE. The navigational performance required solving a route-generating navigational task. The participants were near perfect at developing and transitioning through the new routes in the simple environment (four rooms on each of three floors, with unique identifying objects in each room). This experiment provided an extensive amount of training, three complete guided tours (of all rooms) and an hour of free exploration, which is considerably more than could be provided during mission training or rehearsal on a specific terrain (if the exposure were scaled up for size and complexity). The level of performance could be taken to represent a configurational level of knowledge, although that was not claimed by the authors (Regian, Shebliske, & Monk, 1992).

At least one experiment in our research program has shown the acquisition of an intermediate level of route knowledge in a Virtual building (Witmer, Bailey, & Knerr, 1995) and demonstrated transfer of that knowledge to the real environment. In that experiment subjects were trained to follow a specific route through a building, using verbal instructions and photographs of the building, and through a VE representation of the building. Building-trained participants learned better than VE-trained, who were in turn better than the verbally-trained, when all groups were tested in the actual building. The virtual building used in this research comprised areas on three floors, and included extensive texture mapping detail of office furniture, overhead lights, and distinctive landmarks (e.g., pictures). This large scale space included dozens of offices, work rooms, open cubical bays, and several hallways. Their second experiment used different anchoring (landmarks vs left/right directions) in the instructional strategy, and found that a more exploratory approach tended to improve route learning (Witmer, Bailey, Knerr, & Abel, 1994). Johnson and Wightman (1995) also demonstrated spatial knowledge acquisition and transfer from a VE-based flight regime over a large urban-like airfield terrain to movement and orientation on the actual ground.

Each of these experiments (Johnson & Wightman, 1995; Regian, et al., 1992; Witmer, Bailey, & Knerr, 1995; Witmer, Bailey, Knerr, & Abel, 1994) used a single VE configuration in the acquisition of information, and hence could not compare differences in the VE visual databases (different models). As has been pointed out by previous researchers, landmarks offer unique perceptual patterns that are used in spatial navigation (Lynch, 1960; Seigel, 1981; Siegel & White, 1975). In Seigel's (1981) investigations with children, the boundary or framing conditions aided young children by apparently providing topological positioning cues for isolated landmarks, but did not influence older children in the same fashion. In that experiment, children were exposed to a model of a town which they could then reconstruct using the same model pieces (eliminating problems of verbal or drawing reports of the learned spatial representation). It may be that the older children were developing integrated groups for the "isolated" landmarks by relating those landmarks to others in the spatial array rather than relating them to external framing cues (a high landmark level of knowledge). In accordance with the elaboration and distinctiveness concepts introduced above, people learning a bounded spatial area may use more distinctive landmarks as "anchors" for other landmarks. This may support the construction of subgroups of landmarks which could be revealed through the comparison of spatial areas that differed in landmark distinctiveness. The use of two differing VE databases may provide insight into characteristics that are actually used in constructing mental representations of experienced spaces.

As introduced above, previous research in route learning has shown a passive presentation to be less effective than active movement (Goldin & Thorndyke, 1982). What has not been investigated is whether different VE equipment configurations, which support different levels of interaction with the to-be-learned space, might lead to differences in the level, accuracy, or consistency of spatial knowledge. Further, future implementations of VE equipment for training, briefing, or rehearsal will probably use the most technologically advanced equipment available, under the theory that higher fidelity is better. In fact, distortions of reality (for instructional purposes) may actually become more important in the development of higher level, more coherent mental representations of the to-be-learned area. Whether or not the use of advanced VE configurations will match reality or improve the spatial comprehension is not clear and should not be asssumed. Each of the VE experiments addressed above (Regian, et al., 1992; Witmer, Bailey, Knerr, & Abel, 1994; Witmer, Bailey, & Knerr, 1995; Johnson & Wightman, 1995) used different VE equipment configurations, but did not manipulate equipment or visual database differences within any one experiment. If reality is our current yardstick, then approximating the kinds of interactions that can occur in reality should approximate the kind of learning that occurs in reality. This is again the question of simulation fidelity - what effects do the differences between the simulation and reality have on learning, performance, transfer, and retention. The only way to begin producing a predictive model for these fidelity issues in the VE arena is to examine selected differences in the equipment and presentation materials.

This experiment investigated the potential improvement in spatial knowledge acquisition in two different VE configurations relative to comparable map study. A High-level VE (Hi-VE)

configuration presented a stereoscopic view in a head-mounted display (HMD) linked to head movements (called head-coupling), controlled positional movement through the VE by walking on a treadmill (with handrail buttons for direction control), and used a hand-mounted sensor for pointing at objects in the VE. In this experiment, the Hi-VE was an improvement on the normal VE head-tracked, joystick-controlled configuration, which has shown transfer of VE-acquired spatial knowledge to real building interiors (Witmer, Bailey, & Knerr, 1995). A Low-Level VE (Lo-VE) configuration presented stereographic views in the same HMD with gaze direction and positional movement controlled by a joystick, and used the same hand-mounted sensor for pointing. The Low-VE is akin to fixed-view, joystick directed video-disk or computer-based training configuration, which has been shown in at least one instance to be ineffective for some military spatial orientation activities (Lickteig & Burnside, 1986). A baseline condition using topographical maps was included, as some learning does occur from maps (Goldin & Thorndyke, 1981; 1982), and map briefings are standard fare in U.S. Army training and mission preparation.

Two terrains were also used in the experiment, primarily to ensure that results are not terrain specific (Lynch, 1960). The different terrains allowed a gross test of differences in distinctiveness of landmarks. One of the terrains represents actual terrain that is currently used in U.S. Army training (at Ft. Benning, GA). The other terrain was adapted from a topographical map by adding or enhancing discriminative cues to several features. Participants were required to orient, move through the environment, and perform cognitive terrain appreciation tasks, e.g., identify areas of danger to the participants current location, in order to enhance interaction with and learning of the terrain. The instructional approach was to initially brief the participant on the location of the terrain features relative to the starting position. While interacting with the terrain representation, either in the VE or using maps, participants would be queried about the direction and distance to specific landmarks and provided feedback about the correct direction and distance. The expectation was that the more normal gaze and movement control offered in a Hi-VE configuration should lead to a more accurate or complete spatial representation of the simulated terrains than is found with the more restricted Lo-VE simulation or similar interaction with a topographical map. The timeframe for training was limited given the large area to be learned, and the limited time that participants could commit. The typical participant also had to be trained on topographical representation to a minimum level of comprehension in order to compare learning in the VE with learning on the maps. The end result of the experimental limitations is that the expected general level of spatial knowledge developed would be at the landmark level. The directional, accuracy, and consistency measures (see the results for details) were presumed sensitive enough to discriminate spatial knowledge differences resulting from a single training exercise on the complicated terrains. The comparisons should provide initial indications of VE configuration effects on spatial learning.

Virtual Environments Issues

The learning sessions in this experiment required extensive interaction with paper maps and materials, as well as the VE equipment and the environment databases. Given the varied

population the experimental sample would be drawn from, it seemed reasonable to presume that there would be differences in individual's spatial experiences, and that these differences should be investigated. Therefore, an Individual Differences questionnaire was developed and administered. As there are differences in many cognitive realms and the experiment required remembering two dimensional maps and applying that knowledge to three dimensional areas, a test of cognitive differences in map memory was also applied (Ekstrom, French, Harman, & Dermen, 1976). The individual differences that we believed could influence this investigation are discussed in the next section.

Our previous research has shown quite rapid learning in the performance of simple tasks under varying VE conditions (e.g., Singer, Ehrlich, Cinq-Mars, & Papin, 1995). Unfortunately, previous research has also shown that VE configurations do not support perceptions in the same fashion as the real-world, especially in the area of distance perception (Lampton, McDonald, Singer, & Bliss, 1995; Witmer & Kline, in preparation). Possible problems that might influence spatial acquisition or affect the sensitivity of the performance measures are discussed in the second section, below.

Other research conducted in our program and elsewhere (Lampton, Kolasinski, Knerr, Bliss, Bailey, & Witmer, 1994; Wann, 1993) has indicated that there are frequently sickness problems found in different VE systems that may effect research outcomes. On this basis, simulator sickness is one issue that we regularly include for investigation in experiments within our program, and was investigated during the course of this experiment. The possible effects of simulator sickness and how to measure the phenomenon are briefly introduced in the third section.

Finally, an ongoing focus in our research program is the concept of presence and involvement in VE, how to measure the construct, and what those measurements might mean for learning and performance. Presence in VE systems is briefly addressed in the context of the spatial knowledge research in the last section.

Individual Differences. Given that the use of VE systems is in its infancy, surprising relationships are possible. In order to be thorough, as well as collect possibly relevant background information, several issues were addressed in conjunction with the research. In addition to the normal subject data on age and sex, several questions were asked about health and experience issues (Appendix A). These health questions addressed general fitness, amount of sleep, possible pregnancy, and asked about any history of epilepsy or seizures. Although we did not specifically ask about drug use, one subject that volunteered information about participating in anti-depressant therapy was released. As the experiment was designed to investigate acquisition of spatial knowledge, general questions about participants sense of direction and experience with maps were asked. As we wanted relatively VE-naive participants, questions were asked about previous experience with VE research, games, or exhibitions. These questions were hypothesized to be particularly relevant given the subject recruitment locale (in Orlando,

near Disney World and other attractions that have commercial VE systems). A couple of participants with extensive experience on VE systems or computer-aided design systems were dismissed. Finally, we directly asked about video game and computer experience in order to investigate possible biasing experiences.

People differ in their cognitive abilities, and a visual memory factor has been found in sufficient experiments that three tests for the factor are included in the Kit of Factor-Referenced Cognitive Tests (Ekstrom, French, Harman, & Dermen, 1976). As differential visual memory abilities might confound the investigation of spatial knowledge acquisition, a test for map memory was included in the pre-experimental regime. The obvious concern is that differences in the cognitive ability to remember configuration, location, and orientation might affect the spatial knowledge acquired during a particular experimental experience. The Map Memory test presented portions of twelve maps which were studied for three minutes, followed by an immediate recognition test with five targets embedded in a twelve map array. This sequence was repeated twice in the test.

Distance Perception. One major aspect of spatial knowledge is knowledge of distances. For VE systems this presents a problem in that these systems do not provide all of the cues or the same kind of cues used to support the reasonably accurate estimation of distances in many real world situations. Real-world research investigating egocentric (person-centered) direction and distance in full cue versus reduced cue situations (varying light levels for targets) found that distances shorter than two meters were overestimated and distances over three meters were underestimated (Loomis, Da Silva, Philbeck, & Fukusima, 1996). Research manipulating the amount of visual feedback available in a VE found that the amount of texture available in the visual representation had the largest effect on the static perception of distances (James & Caird, 1995). In research with helicopter pilots using a wide-view helmet-mounted display to present a polygonal database, Wright (1994) found that estimates of forward distances were only 41% of actual; judgments of lateral distances (between objects) were only approximately 50% of actual; and judgments of height were only approximately 72% of actual. The same research found that these same estimations made in the real-world were approximately 90% of actual distances. In research conducted with undergraduates judging shorter distances (less than forty feet) while using a narrow field of view helmet-mounted display and a more textured environment, the distance estimates were also significantly inaccurate (Kline & Witmer, 1996; Lampton, McDonald, Singer, & Bliss, 1995). Other research investigating stereoscopic versus monoscopic presentations in the same environment as Lampton, et al., (1995) has also shown that there is improvement in distance estimation at short distances (less than 30 meters) with stereoscopic presentations that are combined with head coupling (Singer, Ehrlich, Cinq-Mars, & Papin, 1995). The errors committed at very short distances were in the same direction found with judgments of longer distances, with the target being farther away than the estimated distances (Wright, 1994).

The misperception of depth may be an intervening variable in the development of spatial knowledge from VE experiences. The range of presentation distances represented in the

databases (see below) used for this experiment extend from ten to over one thousand meters, which spans the range of distances used in the distance estimation studies mentioned above. Although there is evidence about misperception of distances as short as ten meters (Lampton, McDonald, Singer, & Bliss, 1995; Witmer & Kline, in preparation), as well as at long distance (Wright, 1994) in VE systems, it is not clear that anything can be done about this problem except to use the best available VE display. The state of the technology in presenting visual stimuli is continually changing, and the newest model helmet available was used for this research; the Virtual Research Corporations' VR4^{tm} Helmet-Mounted Display. The VR4 allows adjustments that might alleviate some of the problems in distance perception. For example, the inter-pupilary distance could be set by the user, which would provide more normal cues to the visual system (Rolland, Gibson, & Ariely, 1995). Another reason for using the VR4 in stereoscopic presentation mode was that the typical VE configuration is still stereoscopically based, and we desired to be able to extrapolate the benefits or problems to those configurations. The investigation of visual display presentation differences on distance estimation would present a different kind of experiment, and would almost certainly be rapidly overcome by technological development.

Simulator Sickness. Simulator sickness is a common and constant problem with simulators and training devices (Kennedy, Lane, Berbaum, & Lillenthal, 1993). The symptoms of simulator sickness resemble those of motion sickness (e.g., nausea, headache, stomach awareness, disorientation, sweating, vomiting, fatigue, eyestrain, etc.; Kolasinski, 1995). These symptoms can and often do occur in a simulation even when there is no actual physical movement or motion. The most widely held theory on the origin of motion sickness is the cue conflict theory (see review in Kolasinski, 1995). The cue conflict model proposes that sickness is the result of conflicting information from the visual and the vestibular systems. In other words, the visual system may be registering motion based on the graphics presentation, while the vestibular system senses incongruent or no actual motion. The body is unable to adequately rectify this disparate information, and sickness results. For example, the feeling of vection (illusory self-movement based on visual displays) seems to be a key factor in producing simulator sickness (Hettinger, Berbaum, Kennedy, Dunlap, & Nolan, 1990). Thus, motion and simulator sickness may be different phenomena, potentially with different origins. Whatever the cause, performance decrements may result. For example, Bailey (1994), found that higher post-VR sickness scores correlated with poorer learning of a route through a building.

Kennedy, Lane, Berbaum, and Lillienthal (1993) designed and validated a simulator sickness questionnaire (SSQ) in order to establish differing degress of symptoms related to simulator exposure. They identified items addressing symptoms typically associated with simulator sickness and derived three subscales in addition to a total measure. The first is the nausea scale, which includes the symptoms of general discomfort, increased salivation, sweating, nausea, difficulty concentrating, stomach awareness, and burping. Second is the oculomotor scale, with symptoms addressing general discomfort, fatigue, headache, eye strain, difficulty focusing, difficulty concentrating, and blurred vision. Finally, there is the disorientation scale,

with symptoms covering difficulty focusing, nausea, fullness of head, blurred vision, dizziness with eyes open, dizziness with eyes closed, and vertigo. Each factor or symptom in these subscale is totaled, and the total is weighted to derive the subscale score (with different weights for each subscale; see Kennedy, Lane, Berbaum, & Lillienthal, 1993). The Total simulator sickness score is the total of the totaled symptoms comprising each subscale, with that total multiplied by a different weight. This experiment used the sixteen primary symptoms from the SSQ (Appendix B) to explore the incidence and severity of simulator sickness arising during the experimental regime.

Presence. Presence is the subjective feeling of being immersed in one environment, while actually being physically situated in another. Presence has been hypothesized to be related to or even the basis of improved performance or learning in simulations and VEs (e.g., Sheridan, 1992; Held & Durlach, 1992). A number of factors potentially related to this feeling of presence have been delineated (see Witmer & Singer, 1994, for review). These factors include Control (degree, immediacy, anticipation, mode, and physical or environmental modifiability); Sensory (sensory modality, environmental richness, multi-modal presentation, consistency of multi-modal information, degree of movement perception, and active search); Distraction (isolation, selective attention, and interface awareness), and Realism (scene realism, consistency of information with objective world, meaningfulness of experience, and separation anxiety/disorientation).

Based on these theoretical factors, Witmer and Singer (1994, in preparation) devised and tested several versions of an Immersive Tendencies Questionnaire (ITQ) (Appendix C) and a Presence Questionnaire (PQ) (Appendix D). Several relationships have been tentatively identified between these scales, other measures, and performance in VE (Singer, Witmer, & Bailey, 1994). The most recent versions of these questionnaires (Appendices C & D) were developed based on the analysis of responses to the earlier versions (Witmer & Singer, in preparation). Reliability analyses were used to reduce the scales to items that contributed to reliability, and cluster analyses were performed to determine data-driven subscales (Singer, Witmer, & Bailey, 1994). These analyses identified three clusters in the ITQ that were labeled Involvement, Focus, and Games. The Involvement subscale consists of items that address the respondents tendency to become involved when witnessing something. The Focus subscale items address self-reports of current mental alertness, concentration capabilities, and the ability to block distractions. The Games subscale addresses feeling "inside" video games and frequency of play. The PQ also had three subscales identified: Involved/Control, Natural, and Interface Quality. The Involved/Control subscale addresses the ability to control events in the VE, responsiveness, and visual aspect involvement in the experience. The Natural subscale items address the extent to which interactions feel natural and are consistent with reality. The Interface Quality directly addresses the display and control function interference with the task.

There is some initial evidence for a relationship between presence and performance (Bailey & Witmer, 1994; Lampton, Knerr, Goldberg, Bliss, Moshell, & Blau, 1994). However, that relationship was found with the earlier questionnaire, and different performance tasks. The

positive relationship may arise because presence enables better performance or learning, or the increased performance may lead to a better estimation of involvement, control, etc. Given these possible relationships, the level of presence experienced has become an integral part of the research efforts within the SSRU VIRTUE program. In this experiment presence will be examined using the most recent formulations of the ITQ and PQ, which contain the subscales discussed above.

Method

Subjects

Eighteen females and thirty males were recruited from the University of Central Florida, for a total of forty-eight participants. Participants ranged in age from eighteen to forty-four with a mean of 24.6. Attempts were made to recruit from the local ROTC, with limited success, as they have had introductory training in land navigation and map reading. Participants were required to pass a battery of standard vision tests. A minimum level of ability was required for near acuity, color vision, stereopsis, and the Snellen eye chart (corrective lenses were allowed). Participants passing the eye exam were then administered a fifteen minute training period to introduce them to topographical maps, terrain features, and threat locations. After the training period participants were tested on their topographical map knowledge. Participants not meeting the minimum requirements for topographical map knowledge were excluded from the experiment. All participants were paid $5.00 per hour or given course credit for their participation.

Materials

Questionnaires & Tests. A cognitive test entitled Map Memory (Ekstrom, et al., 1976), was used to test participants ability to recognize or remember previously studied maps. A SSQ (Appendix B) adapted from Kennedy, Lane, Berbaum, & Lillienthal (1993) was also administered. The experiment also used several independently developed questionnaires and tests. One of these was the ITQ (Appendix C) developed by Witmer and Singer (1994). The ITQ is designed to reveal participants tendencies or abilities in focusing on tasks or becoming involved in different activities. The PQ (Appendix D), also developed by Witmer and Singer (1994), was used to obtain subjective reports of the degree of presence experienced and judgments about factors that influenced the experience. Preliminary vision tests included a Snellen eye chart, near-point acuity, Ishihara color perception test, and a test of stereopsis.

Training Materials. A topographical map training packet was created to insure a standard level of topographical map knowledge across groups. The training packet consisted of graphical representations of terrain features (e.g., hills, rivers, saddles) that must be understood in order to use Terrain Association location methods. An explanation of the critical features and attributes of each terrain feature was developed for training, based on standard U.S. Army map training (FM 21-75, 1989). A series of "threat" rules were also presented in conjunction with the

topographical training. The presentation of information was audio-taped in order to standardize the training for all participants. An experimenter was present to provide stock answers if questions were asked during the training.

VE familiarization training was conducted for all experimental conditions using two tasks from the Virtual Environment Performance Assessment Battery (Knerr, Goldberg, et al., 1993). The two tasks were:

> Doorways. An interior movement task requiring controlled movement through ten sequential rooms. The task represents the type and level of difficulty of movement performance required in the VE. The course was formed by a series of 10 rooms connected by a series of 7 x 3 ft. doorways. The position of the doors varied so that a series of non-90 degree turns must be made to navigate the course efficiently. Performance measures included time to cross each room and number of collisions per room.
>
> Fixed Tracking. A pointing task that requires the participant to place a virtual pointing wand on a stationary spherical target The virtual wand was guided by the position of a sensor placed on the first knuckle of the participant's preferred pointing hand. The wand had a grey, semi-opaque cylindrical handle and a line set in the center of one end that was representationally one kilometer long (set for use in the terrains). The target is represented as 0.7 ft. in diameter, and appeared randomly in a three dimensional room between 5 and 19.5 ft. away. The target changed color from red to green when the participants placed the pointing wand on the target (when the line intersected the target). The target disappeared after approximately two seconds of continuous correct pointing. The performance measures included successful completion within the time limit, time of trial, and time to first intercept.

VE Equipment. This experiment was performed at the University of Central Florida, Institute for Simulation and Training, with Visual Systems Laboratory equipment. The visual display information was generated using Performer^{tm} and adjunct specialized software developed by the Institute for Simulation and Training, and was presented using a Silicon Graphics ONYX^{tm}. The visuals were presented through a Virtual Research Corporation VR4 HMD. The VR4 has 48^{O}x36^{O} field of view, with 742x230 color pixels in each lens (Real Time Graphics, 1995). Head and hand motion were tracked by Polhemus Isotrak sensors. A treadmill was instrumented for the Hi-VE condition to enable normal walking speeds, which are translated into a constant walking pace within the terrain databases. The Lo-VE system used the same terrain databases and helmet mounted display (non-head tracked) with movement controlled through the environment by a Gravis 6 degree-of-freedom joystick (which was set to the same constant walking pace as the treadmill). Pointing was accomplished by monitoring a polhemus sensor that was strapped over the first knuckle of the index finger on the preferred hand.

Procedure

The experimental design was a 3x2 factorial design. The VE factor had three levels of experience with the terrain: Hi-VE configuration, Lo-VE configuration, and Map. Two terrains were used in the experiment. Terrain One was an abstract terrain derived from composite topographical maps. Terrain Two was developed by a visual database programmer from a topographical map and aerial photography of an area one kilometer East of the McKenna MOUT site (Military Operations in Urban Terrain) at Ft. Benning. Hi-VE and Lo-VE participants were tested in the terrain in which they practiced, with the Map condition participants using the Hi-VE configuration presentation of the map area that they studied. The few participants with military backgrounds (ROTC students) were assigned evenly to conditions, with one or two in each condition. As unequal numbers of males and females were recruited, proportional assignment was used to distribute equivalent numbers of males and females across conditions and terrains. Because there were development delays on the second terrain, approximately one-half of the total participants for Terrain One completed the experiment before any participants could be assigned to Terrain Two. Within these constraints, the condition and terrain were assigned in a balanced fashion before a participant began, so that the condition and terrain a particular participant experienced was as near random as possible.

Participants were briefed on the nature of the experiment before beginning. After consenting to participate, participants were given the vision tests, completed the SSQ, and completed the ITQ. The SSQ was administered before the training phase of the experiment in order to establish a baseline, and was administered again after the conclusion of the VE test phase. Participants were given topographical training and threat appreciation rules after completing the questionnaires. A recorded presentation (see training materials, above) was made of the set of terrain characteristics for the participants to listen to while they studied the training packet The identification of threat vectors also used taped verbal presentations of the rules while a simple terrain picture was used as an example. The terrain picture portrayed the situation presented in the associated, verbally presented rule. The topographical training sessions lasted approximately fifteen minutes and was the same for all subjects. After training, participants were tested on their topographical and threat identification skills with example topographical maps. These maps tested for identification of terrain features, correct movement direction and distance between locations, identification of threatening areas by application of the threat rules, and measurement of distances between features. A minimum level of topographical knowledge and proficiency with the simple terrain appreciation rules was required for further participation in the research.

A topographical map representation (1cm:100m scale, see Appendix F for a black & white rendition of the terrain maps) of the appropriate terrain was used to brief all participants on the terrain and path to be followed in the experiment. The starting position, intermediate checkpoints, and finishing location were all marked on the topographical map (Appendix F, the connected **x**'s represent the path followed). A verbal presentation identified the marked

positions, named and identified nearby guiding terrain features, and distant relevant guiding terrain points. The verbal presentation also pointed out the directions (in degrees) and distances (in meters) to these landmarks. The participant was then given two minutes after the briefing (which itself lasted approximately two minutes) to study the route, landmarks, and checkpoints.

Following the briefing, participants were trained in their assigned condition on one of the simulated terrains by following the indicated route and identifying requested landmarks at each checkpoint. During the training session, at each of the three checkpoints, participants were asked to locate several of the previously studied landmarks, identify two possible threatening terrain areas, and then cross the terrain following the previously indicated route between checkpoints (Appendix F). The location task was structured so that over the checkpoints, each of the landmarks was queried and identified visually twice. Feedback was provided after landmark identification on the correct orientation and distance to each landmark, and information was provided about the direction and distance to the next checkpoint. Feedback about the current direction and distance to the next checkpoint was also provided during each terrain crossing segment if the participant became lost. No feedback was provided about the correctness of the indicated threat areas, as that task was intended to encourage further cognitive interaction with the terrain. The Map condition participants went through the same sequence of interactions while working with maps of the target terrain. A fresh map was provided at each stage so that previously identified or marked material was not available during the succeeding step. The landmark identification and threat area identification steps were all unavoidably shorter than the same interactions in the VE conditions. The terrain crossing phases (three between the four checkpoints) was timed to approximate the time being taken by the VE subjects to cross the terrain in the VE. The subjects crossing terrain in the VE were not afforded the same kind of study as those in the Map condition, who were instructed to study the terrain and remember the landmarks, compass directions, and distances briefed at each checkpoint. After the training session and before the test, participants were given a five minute break.

The test of spatial knowledge consisted of identifying the landmarks from <u>new</u> positions in the experienced terrain. Participants were "teleported" to these sites (a blackout of the visual display was used to hide the abrupt transitions between the visual displays of the sites) by simply re-rendering the visual display from a new location in the database. At each site the subject was asked to identify (point to) learned landmarks as requested by the experimenter and give a distance estimation to each landmark from each of the six test sites. No feedback was given at any time during the spatial knowledge test.

After completing the test phase participants were again given a post-experiment SSQ, and required to fill out the PQ. Participants were then debriefed, and paid for their allotted time or given class credit. The post VE phase took approximately thirty minutes during which participants were watched for signs of simulator sickness and asked about symptoms indicating simulator sickness. None of the subjects that completed the experiment experienced sufficient distress that more time was required for recovery. The few subjects that could not complete the

experiment were retained until their self-reported symptoms were reduced to the point that they felt quite capable of continuing their normal activities.

Results

Spatial Knowledge Acquisition

There were several ways to consider and analyse the data gathered. The most basic data is whether the participants could correctly point out landmarks from the test sites. Therefore, correct directional indication of landmarks at each site in the VE test was generated. The landmark was scored as correct if the directional indication (vector indicated by pointing) was within the angle subtended by the visible feature from the tested position. The visual angle was used in order to decrease simple pointing error variance. The equipment set precluded rock steady indications of precise landmark centers, although participants were directed to always point out the top center of the landmark. In cases where the angle occupied was very small this angular range was increased slightly, for example, if the feature only occupied one degree the allowable angle was increased to plus or minus two and one-half degrees. This range was selected based on the observed variance in equipment operation for acquiring the directional indication, even with experienced operators. The range allowed for correct identification of non-visible landmarks was plus or minus twenty-two degrees of the center of the feature. This range was selected for two reasons; even the creators of the visual database and experienced operators were often off by as much as ten degrees when indicating non-visible landmarks, and the selected range only provided one chance in eight of being correctly indicated by chance. This allowed the generation of two different but related scales, the number of correctly identified landmarks at each test site (six in each terrain), and the number of times a particular landmark was correctly identified across all test sites. As these measures are somewhat redundant, and any differences found between the test sites would not provide interpretable information about spatial knowledge acquisition, the landmark information was used in the analyses. The landmark totals for each participant were then converted to percentages in three ways; for all directional indications, for only visually available landmarks at the test sites, and for only non-visually available landmarks at the test sites. As the number of times landmarks were either visually or not visually available varied both over test sites and by landmark, a simple count of correct directional indications would have been misleading, and hence the percentages were used.

An Analysis of Covariance (ANCOVA) was performed on the mean overall correct detection percentage (visual and non-visible landmarks averaged together) for each terrain using ITQ Games as covariate. These analyses found significant differences between the individual landmarks identified in each terrain (Terrain One: F=12.53, p<.001; Terrain Two: F=6.55, p<.001). A Post Hoc analysis on the mean percentage correct for each landmark in Terrain One (HSD=18.53, p<.05) indicated that two of the landmarks were identified significantly less frequently than almost all the others (Appendix E). The Post Hoc on landmarks in Terrain Two (HSD=15.45, p<.05) found one landmark that was identified as correct more frequently than the

others (Appendix E). An ANCOVA (again using ITQ Games) on the combined data (averaging over landmarks) also found the experimental groups to be significantly different (F=4.451, p=.018; Hi-VE=.51, Lo-VE=.43, Map=.34), and the terrains were also significantly different (F=4.346, p=.043). A Post Hoc comparison of the differences between the means for the experimental groups found a significant difference only between the Hi-VE and the Map condition (HSD=12.75, p<.05; Difference =16.68). A significant difference between the experimental groups was found with an ANCOVA using only the visually available landmark correct identification percentages (F=6.249, p=.004; Hi-VE=.53, Low-VE=.47, Map=.33), but no significant differences were found with the non-visually available features. A post hoc analysis of the group means for the visually available landmarks also found the Hi-VE significantly different from the Map condition (HSD=18.24, p<.05; Difference=19.32).

Having calculated the percent correct identifications of each landmark in the two terrains, correlations between mean percent correct landmark identifications within each terrain could be examined. The correlation between correctly identified landmarks can serve as an indication of grouping in acquired spatial knowledge. When this correlation matrix was generated for Terrain One, thirteen of the twenty-eight possible correlations (among eight terrain features) were found significant at p<.01 (n=24, Appendix E). A similar matrix generated for Terrain Two (seven terrain features) found only one significant correlation at p<.01 (n=24, Appendix E) out of twenty-one possible correlations. These results seem to indicate that participants were grouping or clustering landmarks, as measured by the correct directional scoring, in Terrain One but not in Terrain Two.

Another measure of the accuracy of participant's spatial knowledge was measured by combining the directional component used previously with a distance estimate, which produces a participant identified point in the three dimensional virtual space for each landmark tested. This technique is referred to as projective convergence (Siegal, 1981). The difference between the actual location and the center of a triangle formed by three of these estimates (taken from different test sites) can be used as a measure of the *accuracy* of the participant's cognitive map. The *consistency* of the participant's mental location for a given landmark is then calculated as the perimeter of the landmark's triangle. (A larger perimeter would indicate a less consistent mental mapping for the landmark.) In this analysis we used only the data from the first three responses for each landmark when it was visible, without screening whether the direction was considered correct. Five landmarks met this criteria in Terrain One, and seven landmarks were visible from at least three sites in Terrain Two. For each landmark, the three localization estimates were recorded as sets of Cartesian coordinates (x,y,z). To calculate a measure of the consistency and accuracy of participant's mental maps for each landmark, three vectors connecting each of the three estimated points for a landmark were derived, forming the borders of a triangle. An average location for each landmark was determined by finding the common point at which bisectors of the three angles of the triangle meet, a measure of the center of the triangle. The distance between this average point and the actual location of the given landmark in the virtual space coordinate system is taken as a measure of *accuracy* of the participant's mental

representation of the landmark location. The three vectors forming the border of the estimated triangle are measured and summed, providing a measure of the *consistency* of the participant's mental representation for that landmark.

A Multiple Analysis of Variance (MANOVA) on the averaged (over landmarks) Accuracy and Consistency measures for the landmarks in the different terrains was performed, as there were different numbers of landmarks contributing to these measures in the different terrains. Both the Accuracy (F=15.95; p<.001) and the Consistency measure (F=10.89; p<.002) were significantly different over the two terrains. Terrain Two was less accurate and more inconsistent over all the landmarks. There was no significant difference between groups on the two measures, nor any interaction between groups and terrains. The analysis of Consistency scores for only Terrain One found a significant difference between experimental conditions (F=11.16, p<.001). A post hoc analysis (HSD=489.19, p=.05) found significant differences between Hi-VE (1143.55) and Low-VE (1704.313; Difference=560.76) and between the Hi-VE and Map conditions (2062.08; Difference=918.45).

The averaged accuracy and consistency scores for the two terrains were examined for relationship with the individual difference questions (Appendix A), the ITQ and PQ (Appendices C & D), and the SSQ data (Appendix B). The only significant correlations found were between each measure and the Interface Quality subscale from the PQ (Accuracy r=-.31, p=.033; Consistency r=-.35, p=.015), and between the Sense of Direction and the Accuracy scores (r=-.32; p=.027.

While not a direct measure of spatial knowledge, the time taken to orient and make directional decisions about landmarks during the test situation provides a partial and indirect measure of confidence in spatial knowledge. When the time taken to complete the test of spatial knowledge was examined, there was a significant difference between the experimental conditions (F=4.582, p=.015). When these times were examined for differences the Hi-VE condition was significantly faster than either the Lo-VE (Difference=314.03) or the Map condition (Difference=491.70, HSD=231.01, p=.05).

Individual Differences

The introductory background questionnaire (Appendix A) addressed a number of potential individual differences that could affect the experimental investigation. The questions focused on health, VE and computer experience, and spatial abilities. The health questions included the Amount of Sleep on the night before the experiment and Motion Sickness Tendency. Spatial abilities questions included Direction Sense, Map Confidence, and Topographical Map Experience. Computer experiences were directly queried in questions about Personal Computer Use, Video-Game Experience, and Virtual Reality Experience. The self-rated attributes of Direction Sense and Map Confidence did not vary significantly by condition or terrain, and the participants did not have significantly different amounts of sleep. There was a positive

correlation between Direction Sense and Map Confidence (r=.36, p=.012). The self-ratings on Motion Sickness Tendency, Personal Computer Use, Topographical Map Experience, Video Game Experience, and Virtual Reality Experience generally had minimal scores (e.g., subjects had little experience with topographical maps, etc.). For those variables, the few experienced participants were relatively evenly distributed over the experimental conditions and terrains. Regression analyses on the entire suite of individual difference variables did not significantly predict any spatial knowledge measures.

There were no statistically significant differences in scores on the topographical feature and threat appreciation post-training test for either experimental condition or terrains. A few participants were washed out of the experiment because they could not pass the topographical feature and threat appreciation test administered after training. Further statistical analyses were not pursued.

Map memory was tested to allow for the possibility that participants with different levels of skill at remembering two-dimensional maps might respond differently in spatial acquisition. An examination of the Map Memory scores showed that there was no difference in the scores on the test across the different experimental conditions, nor in the two terrains (Condition F=.57, p=.57; Terrain F=1.45, p=.54). These results show that there were no imbalances in assignment of different capabilities (at least as measured by this test) to the conditions. Correlations between Map Memory and directional accuracy scores found no significant relationship. Further statistical analyses were not pursued.

Simulator Sickness

The SSQ (Appendix B) was administered both before and after the VE experience. The results of the sixteen item questionnaire were scored according to instructions drawn from Kennedy, Lane, Berbaum, and Lillienthal (1993). A MANOVA on the pre-experiment administration of the SSQ, the post-experiment administration of the SSQ, and the difference scores between pre- and post-administrations did not show any significant differences between experimental groups, terrains, or the interaction of groups and terrains. The t-tests conducted between the pre- and post-experiment scores showed that there were significant differences for all the SSQ scales. The means, t-values, and p-values are presented in Table 1.

Regression analyses did not find any linear or quadratic predictive relationships between simulator sickness and the measures of spatial knowledge (see the section on spatial knowledge acquisition, above). Analyses of the SSQ results that are not directly relevant to spatial acquisition are presented in Appendix B, with a short discussion of those findings.

Table 1

T-Tests on Simulator Sickness Questionnaire Scales Administered Before and After Experimental Sessions

	Mean Pre-Ex.	Mean Post-Ex.	t-value	sig. (2-tail)
NAUSEA	6.36	24.05	-4.57	.001
OCULOMOTOR	9.16	26.57	-5.87	.001
DISORIENTATION	6.67	31.0	-4.83	.001
TOTAL	9.23	30.72	-5.49	.001

Immersive Tendencies & Presence

The ITQ (Appendix C) and PQ (Appendix D) subscales were generated based on scoring instructions from Singer & Witmer (1996, see Appendices C & D). The PQ and ITQ total and subscale means by groups are presented in Table 2 and ancillary analyses of the questionnaires are presented in the respective appendices. When the ITQ scales were investigated for distribution over experimental conditions using a MANOVA (which included the PQ scales, see below), only two scales varied significantly (ITQ Total; $F=3.297$, $p=.047$; ITQ Involvement; $F=3.779$, $p=.031$). An examination of correlations (Pearson's r) between the ITQ Total and subscales and the performance measures found that only the ITQ Games subscale correlated significantly with the mean correct identifications of individual landmarks across sites ($r=.34$, $p=.019$; mean correct directional identifications of visually available landmarks, $r=.31$, $p=.031$; and mean correct directional identifications of non-visually available landmarks, $r=.29$, $p=.047$). [See above for explanation of the spatial knowledge measures.] The difference in the means for ITQ Total and ITQ Involvement is between the Hi-VE and Lo-VE conditions (Table 2), which spans the Map condition, and runs contrary to the spatial knowledge results. The lack of a significant relationship between the subscales and the spatial knowledge measures indicates that the distributional differences in those ITQ subscales did not influence the experiment.

None of the PQ scales differed significantly over the experimental conditions or the terrains, using the MANOVA (see above). The PQ Involvement/Control subscale was significantly correlated with both the mean number of correct landmark directional identifications at test sites ($r=.29$, $p=.044$) and the mean number of correct visually available landmark directional identifications at test sites ($r=.30$, $p=.038$). The PQ Involvement/Control subscale

Table 2

ITQ and PQ means by Experimental Condition.

	High VE	Low VE	Map
ITQ Total*	85.06	72.5	77.69
Focus	37.44	35.56	34.5
Games	6.75	4.88	5.88
Involvement*	30.19	22.44	27.56
PQ Total	95.69	99.06	89.44
Inv/Control	58.19	60.31	55.0
Int.Quality	16.19	16.56	14.88
Naturalness	13.44	13.25	13.56

* p<.05 difference between conditions.

also correlated significantly with mean correct identifications of individual landmarks ($r=.31$, $p=.034$) and mean percent of correctly identified visually available individual landmarks ($r=.34$, $p=.018$). The PQ Interface Quality subscale correlated with the average projective convergency measures (see above for explanation of these spatial knowledge measures) of accuracy ($r=-.31$, $p=.033$) and consistency ($r=-.35$. $p=.015$).

A stepwise linear regression analysis was performed using all of the ITQ and PQ scales with the directional outcome measures. The regression was performed in order to further investigate the relationships revealed through the simple correlation analyses presented above. The regression analysis found significant relations (Multiple $R=.52$; $R^2=.27$) between the PQ Involved/Control subscale ($p=.0029$), the PQ Naturalness subscale ($p=.031$), the ITQ Games subscale ($p=.0434$), and the mean correct directional indication of landmarks at test sites (see above). The inclusion of the ITQ scale in the regression provides further evidence that previous game playing experience eases the acquisition of spatial knowledge in the VE situation, as noted above. The variance absorbed by the PQ scales seems to indicate that these measures address aspects of the VE that can affect spatial learning.

Discussion

Spatial Knowledge Acquisition

The central theme of this experiment was the development of spatial knowledge in two different VE configurations and a map exercise. A measure of directional accuracy for landmarks was used to assess the spatial knowledge acquired during the experimental procedures. Correct directional indication was counted for each landmark at the test sites (counting the number of times each landmark was correctly indicated across all sites). The counts were then converted into percentages based on the number of times landmarks were tested. This measure was also calculated for visually available and non-visually available landmarks. Two other measures of spatial knowledge, referred to as projective convergence measures (Seigal, 1981), are based on the location as indicated by the subject in three dimensional space by a direction and distance estimation from a test site. These measures combine the three dimensional position data from three responses for each tested landmark, providing a measure of the average *accuracy* in the mental model for the particular landmark (the center of the triangle of points for that landmark, see above) and a measure of the *consistency* or variation of the mental representation of the feature (in terms of the perimeter of the triangle of points for that landmark).

The results of analyses using the directional data indicate that a better knowledge of landmarks was acquired in the Hi-VE condition both with overall identification and the identification of only visible landmarks. These data also showed significant differences between the two terrains. The results of the analyses using the accuracy and consistency measures were not as conclusive, showing the Hi-VE to be superior only with the consistency data from Terrain One. Both projective convergence measures showed a significant difference between the terrains, and both measures were negatively correlated with the PQ-Interface Quality. The significant difference between the terrains was probably due to the differences between the distinctiveness of the landmarks in the two terrains. The analysis of Terrain One in isolation supports the differences between the experimental conditions found with the analysis using the directional data. In this case, a difference was also found between the Hi-VE and the Low-VE conditions in addition to the difference found between the Hi-VE and the Map conditions (as with the directional data analyses). Clearly there are differences in the development of spatial knowledge that are based on differences in the learning environment and in the distinctiveness of landmarks.

Better identification of visible landmarks in the Hi-VE indicates the superiority of visual experience-based spatial learning over the symbolic-based spatial learning used in the Map condition. This result may be explained by visual identification factors and the grouping of terrain features. The measures used and analyses applied to those measures limit what this particular single experiment can reveal about the use of visual factors. The VE groups had the advantage of training that enabled learning of visual cues that supported identification of

landmarks, while the Map group had to rely on the encoding of contours and generic vegetation colors for identifying landmarks. This might provide a clear explanation of the results were it not for the low level of VE interaction not being significantly different than the Map condition. This may be related to the difference in interactivity with the environment enabled by the two different VE configurations. The Hi-VE allowed complete head-coupling while the Lo-VE required changing the view by using the joystick. The difference in ease of environment inspection, while not sufficient to produce a significant decrement between Hi-VE and Lo-VE, was sufficient to prevent the Lo-VE configuration from being significantly better than the Map condition. The finding that only the Hi-VE condition was significantly different from Map learning, while the Lo-VE was not significantly different from either, also suggests that there may be something more than just visual recognition behind the results. The significant correlation between the projective convergence measures and PQ-Interface Quality provides some support for this argument. A second and contributing factor may have been the clustering or grouping of landmarks in ways that enhanced the recognition of multiple landmarks from some of the test sites. The positive correlations between landmarks found in Terrain One provides some support for the grouping concept. There may have been something in the physical arrangement between landmarks, especially in terms of certain test sites, that enhanced their recognizability. This would be consistent with the ideas advanced by Siegal and White (1975) about the encoding of urban spaces by older children. This grouping could also be partially or interactively based in the ease of visual scanning, during the training of the Hi-VE group and during testing for both the Hi-VE and Map groups.

As was noted in the introduction, the distinctiveness of the landmarks and their surrounding topographical features may have also contributed to the grouping of some landmarks at some of the test sites and may explain the significant difference between the different terrains. Terrain One, the more abstract terrain, was designed with several more distinctive landmarks (Appendix F). Terrain Two, which replicated the low, rolling terrain at Ft. Benning, did not have the same number or caliber of distinctive features for landmark identification (Appendix F). For example, one of the features from Terrain One that was most frequently identified, and correlated well with other landmarks (Appendix E), was called Twin Hills. These hills were matched, centrally located, and relatively visible from most of the terrain. The hills were also oriented East-West, and thus could provide cues to other landmarks that would contribute to the compass readings. Terrain two was evaluated by subject matter experts and judged to be very representative of the area (SSGT Shonkwiler, personal communication). The subject matter expert also admitted that navigation in the region was to some extent dependent on the use of manmade features (water towers and radio towers) that could be seen at greater distances and provided more refined azimuths for resection. These features were not included in the VE terrain database, nor marked on the topographical maps used in training.

The level of spatial knowledge indicated by these results seems to be at the landmark stage, or slightly better (using Goldin & Thorndyke's 1981 formulation of levels of spatial knowledge). This is the approximate level of spatial knowledge expected in the experiment, based on the

amount of instruction and interaction in the experimental conditions. A better identification of non-visually available landmarks would have indicated a higher level of organization in the mental map constructed by the participants, but this was not found. The significant correlation between the correct identification of visible landmarks and non-visible landmarks indicates that some intermediate level of coherently organized mental maps was being formed, as does the correlations between correctly identified landmarks in Terrain One.

The results indicate that a highly interactive VE experience produces a better mental representation of spatial relationships between landmarks, although the spatial representation differed over the terrains. These results held up when considering the number of landmarks identified at different sites or the number of times a landmark was identified over all sites. For one of the terrains, the consistency of correct positional identification was also significantly better with the more interactive VE over the more typical interactively-constrained VE experience. All of these results indicate the potential for considerable improvement in spatial knowledge through the use of highly interactive virtual environments in training and rehearsal.

In addition, the VE group results and the positive correlations among landmarks indicate that landmarks are not learned individually, but that very early representations are formed that include angles and distances between landmarks or identifying terrain features. The availability of the compass in all conditions could have allowed the participants to learn the offset angles from easily recognizable landmarks, but would have required them to rehearse or learn each of the landmarks angles from each of the other landmarks. This did not seem to be done by the participants, although there was a considerable amount of reference between landmarks. This reference seemed to be in terms of relative position rather than angle offset memorization, as the behavior was looking between landmarks rather than carefully inspecting the compass. This would seem to indicate that humans build spatial knowledge in an incremental fashion by relating landmarks within a spatial area very early in learning. This makes it clear that the three stage descriptions used by some researchers (e.g., Goldin & Thorndyke, 1981; 1982) should not be taken literally as steplike and restricted.

Humans learn from both symbolic information and experience, and learn some things, such as spatial organizations, better through experience than through symbology (Goldin & Thorndyke, 1982). These results make it clear that a VE configuration that allows more normal physical (stereoscopic visual displays) and functional (head-coupled visual displays and walking-based movement) interactions promotes better spatial knowledge acquisition. Given the clear advantage in the early stages of configuration knowledge acquisition shown with short-term exposures to a highly realistic and interactive VE configuration, it is reasonable to infer some benefit for mission training that is based in VE experiences.

Individual Differences

In this experiment, the questions about Sense of Direction, Map use, and Map Confidence were thought to be directly relevant, as was the Map Memory test (Ekstrom, French, Harman, & Dermen, 1976). These pre-experimental measures of possible spatial knowledge related characteristics were analyzed by experimental condition in order to determine whether there was a differential apportionment over the conditions for those variables. The lack of statistical differences in the personal information and the cognitive measures at least shows that none of the conjectured intervening factors were inappropriately distributed over the VE conditions. These data were used to investigate possible contributions to or effect on the spatial acquisition results. Not finding a relationship between the cognitive measure (Map Memory; Ekstrom, et.al., 1976) and spatial knowledge acquisition serves to indicate how little we know about how to measure the cognitive factors that support spatial knowledge acquisition. Certainly the visual aspect of memory would seem to be required, but it is not clear what role the ability to remember and identify map sections has on the ability to remember and identify landmarks in real or virtual spaces.

Simulator Sickness

The significant increase in all subscales of simulator sickness over the course of the experiment is not surprising. The number of participants requesting breaks, or withdrawing from the experiment entirely, has varied between six and sixteen percent in the experiments conducted in our ongoing program (Bailey, 1994; Lampton, Kolasinski, Knerr, Bliss, Bailey, & Witmer, 1994; Witmer, Bailey, and Knerr, 1995). This is not as high as the simulator sickness incidence claimed as an average in other studies (e.g., Regan, 1993). The lack of significant relationship with conditions and performance measures in this experiment is good news from an experimental viewpoint. The lack of significant difference associated with the experimental conditions probably reflects the care taken to equate the stimulus speed and flow in the different movement schemes (the treadmill used in Hi-VE and the joystick used in the Lo-VE). The same care was taken to equate both head-tracking speed and visual presentation frame rates, which varied only in control mechanism between these two conditions. Perhaps most importantly, there was no relationship between the measure of simulator sickness and the measures of spatial knowledge. These results are evidence of our successful efforts to reduce simulator sickness to a minimum in general, and eliminate simulator sickness as a factor influencing the experimental outcomes. The change in simulator sickness over the course of the experiment reflects the general problems in VE in terms of managing update rates, vection during movement, and the distance or perspective cues in the visual display. As noted above, the continued investigation of simulator sickness in research using different VE configurations and manipulating different variables may provide information that can clarify the phenomenon.

Presence

Despite two of the ITQ scales being differentially distributed over the experimental conditions, the construct(s) measured by those scales did not influence the experimental outcome measures. The ITQ Games subscale relationship with the performance measures seems to indicate that more experience with visually presented games makes it easier to develop spatial knowledge from interacting with representational interfaces. There is a need in many games for remembering where entrances or locations are, and this may contribute to some skill in building at least directionally correct spatial knowledge. The data are suggestive at best, and require replication or expansion before firm conclusions can be reached.

The lack of a relationship between VE conditions and presence as measured by the PQ was not a desired outcome. The PQ Total and subscales are designed to measure those elements of the environment that should support the experiencing of presence. However, the significant relationship between the PQ subscales Involved/Control and Naturalness provides some support for the relationship between immersive aspects of VE systems and spatial task performance. The outcomes require further research and examination, and will be dealt with in a later report.

Future Research

The spatial knowledge acquisition investigated in this experiment directly relates to spatial knowledge-based activities performed in standard Infantry ARTEPS. The issues addressed in the acquisition of spatial knowledge are important in developing dismounted soldier simulations for realistic training. The findings reported here contribute to our understanding of how soldier's memory for spatial knowledge is affected by VE experiences. The results also raise issues that should be addressed in order to both further our understanding of spatial knowledge acquisition and evaluate the efficacy of using VE systems for more complicated skill development, mission training, and mission rehearsal. One of the most important issues is whether the superior spatial knowledge from highly interactive VE evidenced in tests in the same system will also result in superior knowledge when transferred to actual open terrain, especially with a soldier population. That experiment has been conducted at Ft. Benning and data analysis has begun. Additional VE efficacy issues are: the rate at which increased training experiences improves spatial acquisition, whether metrics for different terrain configurations can be developed and used to guide training, and improvements in the physical and functional fidelity of the VE systems lead to improvements in spatial knowledge acquisition.

Other related areas of future research have been mentioned above or in the appendices, such as standardizing and continuing the collection of individual information relevant to VE systems and experiences. The ARI, SSRU research program represents a unique opportunity to collect data on responses to many different VE configurations as the VE technology evolves. Investigating background experiences such as computer use, video game experience, and VE game experience will provide relevant information about the younger soldier population

experiences, and raise hypotheses to be investigated. One interesting possiblity indicated by the data gathered in this experiment is that increased experience with VE systems may lead to decreased simulator sickness in future VE experiences. Another interesting finding was that the interface quality of the system (as indicated by the PQ Interface Quality subscale) was related to spatial knowledge acquisition. Further exploration of these and other issues is warranted, and will be continued in ARI SSRU programs.

The SSRU research program has been and continues to be a mixture of basic research, applied research, and proof of concept demonstrations. Future research in spatial knowledge acquisition, team training, and team leader situational awareness will contribute to our understanding of how VE technology can be used to train soldiers for future operations. Along the way it will also develop capabilities for the end goal of an effective demonstration of networked dismounted soldier teams learning and practicing warfighter skills.

References

Bailey, J. P., & Witmer, B. G. (1994, October). Learning and transfer of spatial knowledge in a virtual environment. Proceedings of the 38th Annual Meeting of the Human Factors and Ergonomics Society, Nashville, TN, 1158-1162.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. Journal of Verbal Learning and Verbal Behavior, 11, 671-684.

Craik, F. I. M., & Tulving, E. (1976). Depth of processing and the retention of words in episodic memory. Journal of Experimental Psychology: General, 104, 268-294.

Cross, K. D., & Rugge, S. M. (1982). Map interpretation and terrain analysis course (MITAC) for infantrymen: Illustrated lectures (NPRDC Special Report 82-14). San Deigo, CA: Navy Personnel Research and Development Center.

Department of Defense. (1989). Combat skills of the soldier (FM 21-75). U.S. Government Printing Office.

Ekstrom, R. B., French, J. W., Harman, H. H., & Dermen, D. (1976). Manual for kit of factor-referenced cognitive tests. Princeton. NJ: Educational Testing Service.

Eysenck, M. W., & Keane, M. T. (1990). Cognitive psychology. Hillsdale, NJ: Erlbaum.

Goldin, S. E., & Thorndyke, P. W. (1981, July). Spatial learning and reasoning skill (Rand Corp., R-2805-ARMY, Contract MDA-903-79-C-0549, Santa Monica, CA). Alexandria, VA: U. S. Army Research Institute for the Behavioral and Social Sciences.

Goldin, S. E., & Thorndyke, P. W. (1982). Simulating navigation for spatial knowledge acquisition. Human Factors, 24(4), 457-471.

Hays, R. T., & Singer, M. J. (1989). Simulation fidelity in training system design: Bridging the gap between training and reality. New York: Springer-Verlag.

Held, R. & Durlach, N. (1992). Telepresence. Presence, 1(1), 109-112.

Hettinger, L. J., Berbaum, K. S., Kennedy, R. S., Dunlap, W. P., & Nolan, M. D. (1990). Vection and simulator sickness. Military Psychology, 2(3), 171-181.

Jacobs, R. S., Crooks, W. H., Crooks, J. R., Colburn, E., Fraser, II, R. E., Gorman, P. F., Madden, J. L., Furness, III, T. A., & Tice, S. E. (1994, October). Behavioral Requirements for Training and Rehearsal in Virtual Environments (ARI Technical Report 1011). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A286 311)

James, K. R., & Caird, J. K. (1995). The effects of optic flow, proprioception, and texture on novice locomotion in virtual environments. Proceedings of the Human Factors and Ergonomics 39th Annual Meeting, San Diego, 1405-1409.

Johnson, D. M., & Wightman, D. C. (1995). Using virtual environments for terrain familiarization: Validation (ARI Research Report 1686). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A304 416)

Kennedy, R. S., Lane, N. E., Berbaum, K. S., and Lillenthal, M. G. (1993). A simulator sickness questionnaire (SSQ): A new method for quantifying simulator sickness. International Journal of Aviation Psychology, 3(3), 203-220.

Kirk, R. E. (1968). Experimental design: Procedures for the behavioral sciences. Belmont CA: Wadsworth.

Kline, P. B., & Witmer, B. G. (1996). Distance perception in virtual environments: Effects of field of view and surface texture at near distances. Proceedings of the Human Factors & Ergonomics Society 40th Annual Meeting, .

Knerr, B. W., Lampton, D. L., Bliss, J. P., Goldberg, S. L., Moshel, J. M., & Blau, B. S. (1993, May). Assessing human performance in virtual environments (ARI Working Paper, SOFU 93-01). Orlando FL: U.S. Army Research Institute for the Behavioral and Social Sciences.

Kolasinski, E. M. (1995, May). Simulator sickness in virtual environments (ARI Technical Report 1027). Alexandria, VA: U. S. Army Research Institute for the Behavioral and Social Sciences. (AD A295 861)

Lampton, D. R., Knerr, B. W., Goldberg, S. L., Bliss, J. P., Moshell, J. M., & Blau, B. S. (1994, Spring). The virtual environment performance assessment battery (VEPAB): Development and evaluation. Presence, 3(2).

Lampton, D. R., Kolasinski, E. M., Knerr, B. W., Bliss, J. P., Bailey, J. H., & Witmer, B. G. (1994). Side effects and aftereffects of immersion in virtual environments. Proceedings of the Human Factors and Ergonomics Society 38th Annual Meeting, Nashville, TN, 1154-1157.

Lampton, D. R., McDonald, D. P., Singer, M. J., & Bliss, J. P. (1995). Distance estimation in virtual environments. Proceedings of the Human Factors and Ergonomics Society 39th Annual Meeting, San Diego, CA, 1268-1272.

Lickteig, C. W., & Burnside, B. L., (1986, November). Land Navigation Skills Training: An Evaluation of Computer and Videodisc-Based Courseware (ARI Technical Report 729). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A181 031)

Loomis, J. M., Da Silva, J. A., Philbeck, J. W., & Fukusima, S. S. (1996). Visual perception of location and distance. Current Directions in Psychological Science. 5(3), 72-77.

Lynch, K. (1960). The image of a city. Cambridge, MA: MIT Press.

Paivio, A. (1986). Mental representations: A dual coding approach. In Eysenck & Keane (1990), Cognitive psychology. Hillsdale, NJ: Erlbaum.

Regian, J. W., Shebliske, W. L., & Monk, J. M. (1992). Virtual reality: An instructional medium for visual-spatial tasks. Journal of Communication, 42(4), 136-149.

Rolland, J. P., Gibson, W., & Ariely, D. (1995). Towards quantifying depth and size perception in virtual environments. Presence, 4(1), 24-49.

Sheridan, T. B. (1992). Musings on telepresence and virtual presence. Presence, 1(1), 120-125.

Siegal, A. W. (1981). The externalization of cognitive maps by children and adults: In search of ways to ask better questions. In L.S. Liben, A. H. Patterson, & N. Newcombe (Eds.), Spatial representation and behavior across the life span: Theory and application (pp. 167-194). New York: Academic Press.

Siegal, A. W., & White, S. H. (1975). The development of spatial representations of large-scale environments. In H. Reese (Ed.), Advances in child development and behavior (Vol. 10, pp. 10-55). New York: Academic Press.

Singer, M. J., & Witmer, B. G. (1996). Presence measures for virtual environments: background & scoring instructions (Unpublished manuscript). Orlando, FL: U.S. Army Research Institute for the Behavioral and Social Sciences.

Singer, M. J., Witmer, B. G., & Bailey, J. B. (1994). Development of "presence" measures for virtual environments. Proceedings of the Human Factors & Ergonomics Society 38th Annual Meeting, 983.

Singer, M. J., Ehrlich, J., Cinq-Mars, S., & Papin, J-P. (1995). Task performance in virtual environments: Stereoscopic versus monoscopic displays and head-coupling. Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A306 720)

Wann, J. (1993, September-October). What's wrong with your head mounted display? CyberEdge Journal, Supplemental Issue 17.

Witmer, B. G., Bailey, J. H., & Knerr, B. W. (1995). Training dismounted soldiers in virtual environments: Route learning and transfer (ARI Technical Report 1022). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A292 900)

Witmer, B. G., Bailey, J. H., Knerr, B. W., & Abel, K. (1994). Training dismounted soldiers in virtual environments: Route learning and transfer. Proceedings of the 16th Interservice/Industry Training Systems and Education Conference, Orlando, FL, 2-11.

Witmer, B. G., & Kline, P. B. (in preparation). Training efficiently in virtual environments: Determinants of distance perception of stationary observers viewing stationary objects (ARI Technical Report). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Witmer, B. G., & Singer, M. J. (1994). Measuring immersion in virtual environments (ARI Technical Report 1014). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A286 183)

Witmer, B. G., & Singer, M. J. (in preparation). Measuring presence in virtual environments: A presence questionnaire. Presence.

Wright, R. H. (1994). Virtual reality psychophysics: Forward and lateral distance, height, and speed perceptions with a wide angle helmet display (ARI Technical Report 1025). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A294 027)

Appendix A

Personal Data Questionnaire

SUBJECT NUMBER: __________ DATE: _________

BIOGRAPHICAL DATA FORM

Please fill in the blank or circle the correct or most appropriate answer. THIS INFORMATION IS USED FOR BACKGROUND INFORMATION ONLY AND WILL BE HELD PRIVATE. IT WILL ONLY BE USED IN COMBINATION WITH DATA FROM OTHER SUBJECTS.

1. Age ________

2. Gender: M F

3. Do you have a history of epilepsy or seizures? YES NO

4. Are you in your usual state of fitness? YES NO

5. How many hours of sleep did you get last night? _____ hours

6. Was the amount of sleep sufficient? YES NO

7. Is there a possibility you are pregnant? YES NO

8. Do you have a good sense of direction?

1	2	3	4	5
POOR		AVERAGE		VERY GOOD

9. Have you ever had motion sickness? YES NO
   (As with nausea, blurred vision, headaches, etc.)

   If yes: How susceptible do you feel you are?

1	2	3	4	5	6	7
very mildly (seldom)			average (occasionally)			very highly (often)

9. Have you ever participated in any kind of Virtual Reality research or testing? YES NO
[VR means wearing a 3D helmet, using a hand control, and/or wearing a glove.]

10. Have you ever experienced a Virtual Reality Game or Entertainment? YES NO

If you answered yes, how many times in the last year have you experienced a VR game or entertainment?

1 2 3 4 5 6 7 8 9 10 11+

11. How many hours per week do you play video games (either home or arcade)? ________

12. How many hours per week do you use computers? __________

13. How good are you at using road or street maps?

1	2	3	4	5
POOR		AVERAGE		VERY GOOD

14. Do you have any experience or training with topographical maps? YES NO

If you answered yes, how many times in the last year have you used a topographical map?

1 2 3 4 5 6 7 8 9 10 11+

Appendix B

Simulator Sickness Questionnaire

Subject Number __________
Date __________

COMFORT QUESTIONNAIRE

Instrutions: Please indicate the severity of symptoms that apply to you right now.

1. General Discomfort	None	Slight	Moderate	Severe
2. Fatigue	None	Slight	Moderate	Severe
3. Headache	None	Slight	Moderate	Severe
4. Eye Strain	None	Slight	Moderate	Severe
5. Difficulty Focusing	None	Slight	Moderate	Severe
6. Increased Salivation	None	Slight	Moderate	Severe
7. Sweating	None	Slight	Moderate	Severe
8. Nausea	None	Slight	Moderate	Severe
9. Difficulty Concentrating	None	Slight	Moderate	Severe
10. Fullness of Head	None	Slight	Moderate	Severe
11. Blurred Vision	None	Slight	Moderate	Severe
12. Dizzy (Eyes Open)	None	Slight	Moderate	Severe
13. Dizzy (Eyes Closed)	None	Slight	Moderate	Severe
14. Vertigo*	None	Slight	Moderate	Severe
15. Stomach Awareness**	None	Slight	Moderate	Severe
16. Burping	None	Slight	Moderate	Severe

*Vertigo is a disordered state in which the person or his/her surroundings seem to whirl dizzily: giddiness.

**Stomach awareness is usally used to indicate a feeling of discomfort which is just short of nausea.

Are there any other symptoms you are experiencing right now? If so, please describe the symptom(s) and rate its/their severity below. Use the other side if necessary.

Simulator Sickness Questionnaire Results

The first administration of the SSQ measure showed a correlation with several of the self-rated individual difference questions (Appendix A). Participants Sense-of-Direction scores were negatively correlated with the pre-experiment Disorientation subscale (r=-.41, p=.004), the Oculomotor subscale (r=-.47, p=.001), and the overall SSQ Total (r=-.41, p=.004). As might be expected from the significant correlation with Sense-of-Direction, Confidence-with-Maps also correlated negatively with the pre-experiment administration of the Oculomotor subscale (r=-.32. p=.026) and the SSQ Total (r=-.32, p=.026). The relationship indicates that participants with fewer symptoms before the experiment began also reported having a better sense of direction and higher confidence with maps while those with more symptoms reported having a poor sense of direction and low confidence with maps.

The post-experiment administration of the SSQ also correlated significantly with two of the pre-experiment individual difference questions. Responses on Tendency-to-Motion-Sickness (N=16 positive responses; scaled one [seldom] to five [often]) were significantly positively correlated with the post-experiment Disorientation subscale (r=.51, p=.044) and SSQ Total (r=.50, p=.048). The responses from the nine participants who had some prior experience with Virtual Reality were significantly negatively correlated with the post-experiment Nausea subscale (r=-.68, p=.042). The VR-experience responses also showed a significant negative correlation with the change in three SSQ scales over the course of the experiment (where the change is calculated as post-score minus pre-score). VR-Experience correlated with the Disorientation difference (r=-.80, p=.01), Oculomotor difference (r=-.74, p=.023) and SSQ Total difference (r=-.80, p=.01). This seems to indicate that more experience with VR systems leads to less change in symptomology over the course of an experience.

Almost all of the post-experiment SSQ and the SSQ difference scores were also significantly negatively correlated with the scores on the topographical test (which tested acquisition of topographical information after training). On the post-experiment administration these were the Nausea subscale (r=-.35, p=.014), the Oculomotor subscale (r=-.29, p=.044), and the SSQ total (r=-.33, p=.02). For the difference scores, all of the subscales (Nausea r=-.32, p=.029; Disorientation r=-.33, p=.022; Oculomotor r=-.34, p=.017; Total r=-.34, p=.017) were negative and significant. This indicates that participants that did well on the training did not develop higher levels of simulator sickness symptoms while those that did more poorly developed more or higher levels of symptoms. Finally, neither the post-experiment nor the difference scores were significantly related to the time spent in the VE performing the training and testing.

Discussion of Simulator Sickness Questionnaire Results

Finding that subscales from the initial administration of the validated SSQ (Kennedy, et al., 1993) correlated negatively with self-ratings of Sense-of-Direction and Map-Confidence is interesting, but minimally informative as neither the post-experimental administration nor the SSQ difference scores had the same relationship. If there had been a relationship between these or other individual difference factors and the post-experience or difference scores, perhaps some new questions would have been raised about the basis of simulator sickness. Finding a positive relationship between Tendency-to-Motion-Sickness self-ratings and SSQ responses merely indicates reasonable self-knowledge by the participants. More interesting is the relationship between VE-experience and SSQ results. These findings (although with a low number of participants) indicates a possible adaptation to VE experiences.

The relationship found with the training test and difference subscales might indicate a relationship with performance, but there were no other relationships found with spatial knowledge performance. The lack of relationship with the amount of time spent in the VE is probably significant, as previous work has shown time to be an important determinant of both changes and levels of symptomology.

The incidence of simulator sickness remains a cause of concern for the general implementation of VE systems (Kennedy & Stanney, 1996). As the use of VE systems for testing, evaluation, and training increases, this issue will become even more important. The results of this experiment did not directly address a particular simulator sickness factor or relationship, but suggest that the amount of time spent in the faster updating VE systems currently available is not as debilitating as the older systems (Lampton, Kolasinski, Knerr, Bliss, Bailey, & Witmer, 1994).

Reference List for Appendix.

Kennedy, R. S., & Stanney, K. M. (1996). Postural instability induced by virtual reality exposure: Development of a certification protocol. International Journal of Human-Computer Interaction, 8(1), 25-47.

Appendix C

Immersive Tendencies Questionnaire

IMMERSIVE TENDENCIES QUESTIONNAIRE
(Witmer & Singer, Version 3.0, Feb. 1995)

Indicate your preferred answer by marking an "X" in the appropriate box of the seven point scale. Please consider the entire scale when making your responses, as the intermediate levels may apply. For example, if your response is once or twice, the second box from the left should be marked. If your response is many times but not extremely often, then the sixth (or second box from the right) should be marked.

1. Do you easily become deeply involved in movies or tv dramas?

|_______|_______|_______|_______|_______|_______|_______|

NEVER OCCASIONALLY OFTEN

2. Do you ever become so involved in a television program or book that people have problems getting your attention?

|_______|_______|_______|_______|_______|_______|_______|

NEVER OCCASIONALLY OFTEN

3. How mentally alert do you feel at the present time?

|_______|_______|_______|_______|_______|_______|_______|

NOT ALERT MODERATELY FULLY ALERT

4. Do you ever become so involved in a movie that you are not aware of things happening around you?

|_______|_______|_______|_______|_______|_______|_______|

NEVER OCCASIONALLY OFTEN

5. How frequently do you find yourself closely identifying with the characters in a story line?

|_______|_______|_______|_______|_______|_______|_______|

NEVER OCCASIONALLY OFTEN

6. Do you ever become so involved in a video game that it is as if you are inside the game rather than moving a joystick and watching the screen?

|_______|_______|_______|_______|_______|_______|_______|

NEVER OCCASIONALLY OFTEN

7. What kind of books do you read most frequently? (CIRCLE ONE ITEM ONLY!)

Spy novels	Fantasies	Science fiction
Adventure novels	Romance novels	Historical novels
Westerns	Mysteries	Other fiction
Biographies	Autobiographies	Other non-fiction

8. How physically fit do you feel today?

|_________|_________|_________|_________|_________|_________|_________|

NOT FIT — MODERATELY FIT — EXTREMELY FIT

9. How good are you at blocking out external distractions when you are involved in something?

|_________|_________|_________|_________|_________|_________|_________|

NOT VERY GOOD — SOMEWHAT GOOD — VERY GOOD

10. When watching sports, do you ever become so involved in the game that you react as if you were one of the players?

|_________|_________|_________|_________|_________|_________|_________|

NEVER — OCCASIONALLY — OFTEN

11. Do you ever become so involved in a daydream that you are not aware of things happening around you?

|_________|_________|_________|_________|_________|_________|_________|

NEVER — OCCASIONALLY — OFTEN

12. Do you ever have dreams that are so real that you feel disoriented when you awake?

|_________|_________|_________|_________|_________|_________|_________|

NEVER — OCCASIONALLY — OFTEN

13. When playing sports, do you become so involved in the game that you lose track of time?

|_____|_____|_____|_____|_____|_____|_____|

NEVER OCCASIONALLY OFTEN

14. How well do you concentrate on enjoyable activities?

|_____|_____|_____|_____|_____|_____|_____|

NOT AT ALL MODERATELY WELL VERY WELL

15. How often do you play arcade or video games? (OFTEN should be taken to mean every day or every two days, on average.)

|_____|_____|_____|_____|_____|_____|_____|

NEVER OCCASIONALLY OFTEN

16. Have you ever gotten excited during a chase or fight scene on TV or in the movies?

|_____|_____|_____|_____|_____|_____|_____|

NEVER OCCASIONALLY OFTEN

17. Have you ever gotten scared by something happening on a TV show or in a movie?

|_____|_____|_____|_____|_____|_____|_____|

NEVER OCCASIONALLY OFTEN

18. Have you ever remained apprehensive or fearful long after watching a scary movie?

|_____|_____|_____|_____|_____|_____|_____|

NEVER OCCASIONALLY OFTEN

19. Do you ever become so involved in doing something that you lose all track of time?

|_____|_____|_____|_____|_____|_____|_____|

NEVER OCCASIONALLY OFTEN

20. On average, how many books do you read for enjoyment in a month?

|_____|_____|_____|_____|_____|_____|_____|

NONE ONE TWO THREE FOUR FIVE MORE

21. Do you ever get involved in projects or tasks, to the exclusion of other activities?

|__________|__________|__________|__________|__________|__________|__________|

NEVER OCCASIONALLY OFTEN

22. How easily can you switch attention from the activity in which you are currently involved to a new and completely different activity?

|__________|__________|__________|__________|__________|__________|__________|

NOT SO EASILY FAIRLY EASILY QUITE EASILY

23. How often do you try new restaurants or new foods when presented with the opportunity?

|__________|__________|__________|__________|__________|__________|__________|

NEVER OCCASIONALLY FREQUENTLY

24. How frequently do you volunteer to serve on committees, planning groups, or other civic or social groups?

|__________|__________|__________|__________|__________|__________|__________|

NEVER SOMETIMES FREQUENTLY

25. How often do you try new things or seek out new experiences?

|__________|__________|__________|__________|__________|__________|__________|

NEVER OCCASIONALLY OFTEN

26. Given the opportunity, would you travel to a country with a different culture and a different language?

|__________|__________|__________|__________|__________|__________|__________|

NEVER MAYBE ABSOLUTELY

27. Do you go on carnival rides or participate in other leisure activities (horse back riding, bungee jumping, snow skiing, water sports) for the excitment of thrills that they provide?

|__________|__________|__________|__________|__________|__________|__________|

NEVER OCCASIONALLY OFTEN

28. How well do you concentrate on disagreeable tasks?

|______|______|______|______|______|______|______|

NOT AT ALL — MODERATELY WELL — VERY WELL

29. How often do you play games on computers?

|______|______|______|______|______|______|______|

NOT AT ALL — OCCASIONALLY — FREQUENTLY

30. How many different video, computer, or arcade games have you become reasonably good at playing?

|______|______|______|______|______|______|______|

NONE — ONE — TWO — THREE — FOUR — FIVE — MORE

31. Have you ever felt completely caught up in an experience, aware of everything going on and completely open to all of it?

|______|______|______|______|______|______|______|

NEVER — OCCASIONALLY — FREQUENTLY

32. Have you ever felt completely focused on something, so wrapped up in that one activity that nothing could distract you?

|______|______|______|______|______|______|______|

NOT AT ALL — OCCASIONALLY — FREQUENTLY

33. How frequently do you get emotionally involved (angry, sad, or happy) in news stories that you see, read, or hear?

|______|______|______|______|______|______|______|

NEVER — OCCASIONALLY — OFTEN

34. Are you easily disturbed when involved in an activity or working on a task?

|______|______|______|______|______|______|______|

NEVER — OCCASIONALLY — OFTEN

Scoring Instructions

Simply score the boxes for each question from left to right beginning with one and increasing in value to the box the subject has marked, and the number of that box becomes the score. The subscale scores are the sum of the scores for each subscale item. There is no weighting of items or subscales. The questionnaire total and subscales are comprised as follows:

IMMERSIVE TENDENCIES QUESTIONNAIRE

<u>Total</u>: Items 1, 2, 3, 4, 5, 6, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19.
<u>ITQ-Focus</u>: Items 1, 3, 8, 9, 13, 16, & 19.
<u>ITQ-Involvement</u>: Items 2, 4, 5, 11, 12, 17, & 18.
<u>ITQ-Games</u>: Items 6 & 15.

New questions have been added to the questionnaire, but should not be added to the total or subscales as they are just beginning to be investigated. The new (unanalyzed) questions are scored the same as the other questions. None of the new questions seem to require reverse scoring.

Immersive Tendencies Questionnaire Results

A correlation analysis of the two questionnaires (ITQ and PQ) found some of the subscales in these two questionnaires to be significantly correlated. The ITQ Focus subscale correlated significantly with PQ Involved/Control (Pearson r=.36, p=.013), PQ Naturalness (r=.34, p=.018), and PQ Total (r=.34, p=.017). The ITQ Involvement subscale was negatively correlated with the PQ Interface Quality subscale (r=-.31, p=.031).

The only ITQ subscale that correlated with the individual data questions was the Games subscale, which correlated with the video games experience question (r=.30, p=.036). There were only two significant correlations between the ITQ and the pre-experiment SSQ scales; ITQ-Focus and Disorientation correlated negatively (r=-.33, p=.022) while ITQ-Involvement and Oculomotor Discomfort correlated positively (r=.34, p=.018). There were no correlations between the ITQ subscales and the post-experiment SSQ scales, nor with the difference between pre and post experiment SSQ scales.

In this experiment only four of the possible sixteen correlations between the ITQ and the PQ scales were significant. ITQ Focus correlated with PQ Involved/Control (r=.36, p=.013), with PQ Natural (r=.34; p=.018), and with PQ Total (r=.34; p=.017). ITQ Involvement correlated significantly with PQ Interface Quality (r=-.31; p=.031). These correlations are fewer than were anticipated when the questionnaires were last revised (Witmer & Singer, in preparation).

Discussion of ITQ Results

The significant negative correlation found between Involvement and PQ Interface Quality, when positive correlations are expected, indicates that those who rated their Involvement capability high also rated the interfaces used in the experiment as low and those that rated their Involvement capability low rated the interfaces more highly. The significant correlations between the ITQ Focus subscale and the PQ subscales Involved/Control, Naturalness, and Total is reasonable, given that the scales are addressing related issues (Witmer & Singer, 1994). The ITQ-Focus subscale addresses the participant's self-rated tendency to be able to focus on the tasks at hand, which is theoretically linked to experiencing of presence.

The ITQ-Games correlation with the individual data question on Video games experience is a result of the overlap of issues, as video game experience is one of the Games subscale questions. The reason for the correlations between ITQ subscales and the SSQ pre-experiment subscales is not clear.

The significant difference in ITQ Total and Involvement over the experimental conditions and the lack of relationship with the experimental outcomes raises some problems for the conceptual structure being used for the questionnaire. These scales were both lowest within the Low-VE group and highest within the Hi-VE group. Based on the reasoning used in constructing the questionnaires, this should have amplified the differences found between the experimental conditions, and the outcomes for the PQ. However, the Involvement subscale is focused on engagement in passive activities, and so may not have had any effect on the active interaction with the environment required in the Low-VE condition.

Appendix D

Presence Questionnaire

PRESENCE QUESTIONNAIRE
(Witmer & Singer, Versus 3.0, Nov. 1994)

Characterize your experience in the environment, by marking an "X" in the appropriate box of the 7-point scale, in accordance with the question content and descriptive labels. Please consider the entire scale when making your responses, as the intermediate levels may apply. Answer the questions independently in the order that they appear. Do not skip questions or return to a previous question to change your answer.

WITH REGARD TO THE EXPERIENCED ENVIRONMENT

1. How much were you able to control events?

|_________|_________|_________|_________|_________|_________|_________|

NOT AT ALL SOMEWHAT COMPLETELY

2. How responsive was the environment to actions that you initiated (or performed)?

|_________|_________|_________|_________|_________|_________|_________|

NOT RESPONSIVE MODERATELY RESPONSIVE COMPLETELY RESPONSIVE

3. How natural did your interactions with the environment seem?

|_________|_________|_________|_________|_________|_________|_________|

EXTREMELY ARTIFICIAL BORDERLINE COMPLETELY NATURAL

4. How much did the visual aspects of the environment involve you?

|_________|_________|_________|_________|_________|_________|_________|

NOT AT ALL SOMEWHAT COMPLETELY

5. How much did the auditory aspects of the environment involve you?

|_________|_________|_________|_________|_________|_________|_________|

NOT AT ALL SOMEWHAT COMPLETELY

6. How natural was the mechanism which controlled movement through the environment?

|__________|__________|__________|__________|__________|__________|__________|

EXTREMELY ARTIFICIAL	BORDERLINE	COMPLETELY NATURAL

7. How compelling was your sense of objects moving through space?

|__________|__________|__________|__________|__________|__________|__________|

NOT AT ALL	MODERATELY COMPELLING	VERY COMPELLING

8. How much did your experiences in the virtual environment seem consistent with your real world experiences?

|__________|__________|__________|__________|__________|__________|__________|

NOT CONSISTENT	MODERATELY CONSISTENT	VERY CONSISTENT

9. Were you able to anticipate what would happen next in response to the actions that you performed?

|__________|__________|__________|__________|__________|__________|__________|

NOT AT ALL	SOMEWHAT	COMPLETELY

10. How completely were you able to actively survey or search the environment using vision?

|__________|__________|__________|__________|__________|__________|__________|

NOT AT ALL	SOMEWHAT	COMPLETELY

11. How well could you identify sounds?

|__________|__________|__________|__________|__________|__________|__________|

NOT AT ALL	SOMEWHAT	COMPLETELY

12. How well could you localize sounds?

|__________|__________|__________|__________|__________|__________|__________|

NOT AT ALL	SOMEWHAT	COMPLETELY

13. How well could you actively survey or search the virtual environment using touch?

|_________|_________|_________|_________|_________|_________|_________|

NOT AT ALL SOMEWHAT COMPLETELY

14. How compelling was your sense of moving around inside the virtual environment?

|_________|_________|_________|_________|_________|_________|_________|

NOT COMPELLING MODERATELY COMPELLING VERY COMPELLING

15. How closely were you able to examine objects?

|_________|_________|_________|_________|_________|_________|_________|

NOT AT ALL PRETTY CLOSELY VERY CLOSELY

16. How well could you examine objects from multiple viewpoints?

|_________|_________|_________|_________|_________|_________|_________|

NOT AT ALL SOMEWHAT EXTENSIVELY

17. How well could you move or manipulate objects in the virtual environment?

|_________|_________|_________|_________|_________|_________|_________|

NOT AT ALL SOMEWHAT EXTENSIVELY

18. How involved were you in the virtual environment experience?

|_________|_________|_________|_________|_________|_________|_________|

NOT AT ALL INVOLVED MILDLY INVOLVED COMPLETELY ENGROSSED

19. How much delay did you experience between your actions and expected outcomes?

|_________|_________|_________|_________|_________|_________|_________|

NO DELAYS MODERATE DELAYS LONG DELAYS

20. How quickly did you adjust to the virtual environment experience?

|_________|_________|_________|_________|_________|_________|_________|

NOT AT ALL SLOWLY LESS THAN

ONE MINUTE

21. How proficient in moving and interacting with the virtual environment did you feel at the end of the experience?

|________|________|________|________|________|________|________|

NOT PROFICIENT	REASONABLY PROFICIENT	VERY PROFICIENT

22. How much did the visual display quality interfere or distract you from performing assigned tasks or required activities?

|________|________|________|________|________|________|________|

NOT AT ALL	INTERFERED SOMEWHAT	PREVENTED TASK PERFORMANCE

23. How much did the control devices interfere with the performance of assigned tasks or with other activities?

|________|________|________|________|________|________|________|

NOT AT ALL	INTERFERED SOMEWHAT	INTERFERED GREATLY

24. How well could you concentrate on the assigned tasks or required activities rather than on the mechanisms used to perform those tasks or activities?

|________|________|________|________|________|________|________|

NOT AT ALL	SOMEWHAT	COMPLETELY

25. How completely were your senses engaged in this experience?

|________|________|________|________|________|________|________|

NOT AT ALL ENGAGED	MILDLY ENGAGED	COMPLETELY ENGAGED

26. To what extent did events occurring outside the virtual environment distract from your experience in the virtual environment?

|________|________|________|________|________|________|________|

NOT AT ALL	MODERATELY	VERY MUCH

27. Overall, how much did you focus on using the display and control devices instead of the virtual experience and experimental tasks?

|_________|_________|_________|_________|_________|_________|_________|

NOT AT ALL SOMEWHAT VERY MUCH

28. Were you involved in the experimental task to the extent that you lost track of time?

|_________|_________|_________|_________|_________|_________|_________|

NOT AT ALL SOMEWHAT COMPLETELY

29. How easy was it to identify objects through physical interaction; like touching an object, walking over a surface, or bumping into a wall or object?

|_________|_________|_________|_________|_________|_________|_________|

IMPOSSIBLE MODERATELY DIFFICULT VERY EASY

30. Were there moments during the virtual environment experience when you felt completely focused on the task or environment?

|_________|_________|_________|_________|_________|_________|_________|

NONE OCCASIONALLY FREQUENTLY

31. How easily did you adjust to the control devices used to interact with the virtual environment?

|_________|_________|_________|_________|_________|_________|_________|

DIFFICULT MODERATE EASILY

32. Was the information provided through different senses in the virtual environment (e.g., vision, hearing, touch) consistent?

|_________|_________|_________|_________|_________|_________|_________|

NOT CONSISTENT SOMEWHAT CONSISTENT VERY CONSISTENT

Scoring Instructions

Simply score the boxes for each question from left to right beginning with one and increasing in value to the box the subject has marked, and the number of that box becomes the score. Some of the questions have reversed response anchors, and are scored so the left-most box receives a seven and the rest decrease in value. The subscale scores are the sum of the scores for each subscale item. There is no weighting of items or subscales. The questionnaire total and subscales are comprised as follows:

PRESENCE QUESTIONNAIRE

<u>Total</u>: Items 1, 2, 3, 4, 6, 7, 8, 9, 10, 14, 15, 16, 18, 19+, 20, 21, 22+, 23+, 24.
<u>PQ-Involved/Control</u>: Items 1, 2, 4, 7, 9, 10, 14, 18, 19+, 20, & 21.
<u>PQ-Natural</u>: Items 3, 6, & 8.
<u>PQ-Interface Quality</u>: Items 22+, 23+, & 24.
<u>PQ-Auditory*</u>: Items 5, 11, 12.
<u>PQ-Haptic*</u>: Items 13 & 17.
<u>PQ-Resolution*</u>: Items 15 & 16.

The last three subscales listed for the PQ are marked with an asterisk (*) because they have yet to be used in analyses, but are being retained on a theoretical basis. Since there have been no haptic or auditory interfaces, nor any differences in resolution to judge, those items have been scored as zero. Items marked with a plus (+) have to be reverse scored (see above) in order to contribute to the subscale and overall totals.

New questions have been added to the questionnaire, but should not be added to the total or subscales as they are just beginning to be investigated. The new (unanalyzed) questions are scored the same as the other questions. None of the new questions seem to require reverse scoring.

Presence Questionnaire Results

The PQ subscale Interface Quality correlated significantly with the Map Confidence question ($r=.35$, $\underline{p}=.014$) and the self-rating for Motion Sickness ($r=-.36$, $\underline{p}=.012$). There were also significant negative correlations between the PQ-Interface Quality scale and the post-experiment SSQ Disorientation ($r=-.34$, $\underline{p}=.019$), Oculomotor ($r=-.33$, $\underline{p}=.021$), and Total scales ($r=-.36$, $\underline{p}=.02$). The correlations between the SSQ Difference scores and the PQ were all negative but non-significant. Finally, the PQ-Involvement/Control scale correlated positively with the Total Time spent in the VE ($r=.32$, $\underline{p}=.028$). The PQ Total also correlated positively with the Total Time spent in the VE ($r=.34$, $\underline{p}=.017$) and with the Training Time in VE ($r=.31$, $\underline{p}=.031$).

Discussion

The data from this experiment indicates that efforts to reduce simulator sickness while enhancing presence (through improved interface capabilities) does not seem to cancel the generally negative, but non-significant, relationship between sickness and presence. The positive correlation of presence with time in the VE (the first time we have examined this relationship) suggests that the more time spent in the VE without getting sick, the more normal the experience seems to be. Previous analyses have indicated that in general the more time one spends in a VE the greater the incidence of simulator sickness (Lampton, Kolasinski, Knerr, Bliss, Bailey, & Witmer, 1994). There may be an adaptation that is occuring, which would suggest that presence should continue to increase with more normal interface configurations and more time in the VE. This may also present problems if the adaptation that is occuring conflicts with normal real world adaptations. The PQ Total and subscales are designed to measure those elements of the environment that should support the experiencing of presence. The outcomes require further research and examination, and will be dealt with in a later report. This is contrary to the result that the authors of the scale anticipated (Witmer & Singer, 1994). The relationship between the ITQ and the PQ was supposed to be consistently positive, and in general had been positive in previous research (Witmer & Singer, 1994). The negative relationship between self-reports of involvement (ITQ Involvement) and the judged interface quality of the VE configurations (PQ Interface Quality) are confusing and will be investigated further when the data from this experiment are combined with questionnaire data from other experiments.

This is contrary to the result that the authors of the scale anticipated (Witmer & Singer, 1994). The relationship between the ITQ and the PQ was supposed to be consistently positive, and in general had been positive in previous research (Witmer & Singer, 1994). The negative relationship between self-reports of involvement (ITQ Involvement) and the judged interface quality of the VE configurations (PQ Interface Quality) are confusing and will be investigated further when the data from this experiment are combined with questionnaire data from other experiments.

Appendix E

Analyses of Landmarks Within Terrains

TERRAIN ONE:

Table of Differences between percent correct identification of Landmarks over sites, including mean percent correct for each Landmark. Column Mean is subtracted from the Row Mean to obtain the Difference Score. A negative value equates to better identification of the Landmark identified by the column number than the Landmark identified in the row.

MEAN	2	3	4	5	6	7	8
1: 28.13	01.05	-13.54	-31.59*	-39.23*	-30.2*	-25.*	-20.48*
2: 27.08		-14.59	-32.64*	-40.28*	-31.25*	-26.08*	-21.53*
3: 41.67			-18.05	-25.69*	-16.66	-11.46	-06.94
4: 59.72				-07.64	01.39	06.59	11.11
5: 67.36					09.03	14.23	18.75*
6: 58.33						00.52	09.72
7: 53.13							04.52
8: 48.61							

* p<.05 [HSD = 18.53]

Table of Correlations between Landmarks in Terrain One (N=24).

	1: Bob's	2	3	4	5	6	7
2: Mentor	.4704						
	p=.07						
3: Center	.5703	.1580					
	p=.02	p=.461					
4: Kevin's	.5730	.2174	6495				
	p=.003	p=.307	p=.001				
5: Twins	.6180	.1754	.5779	.5719			
	p=.001	p=.412	p=.003	p=.004			
6: Rd Cut	.3878	.3239	.5205	.6030	.6702		
	p=.061	p=.123	p=.009	p=.002	p<.001		
7: Rd Int	.4512	.2738	.2823	.4936	.3861	.4598	
	p=.027	p=.195	p=.181	p=.014	p=.062	p=.024	
8: Abel	.5233	.1528	.5215	.7412	.6537	.5296	.2378
	p=.009	p=.476	p=.009	p<.001	p<.001	p=.008	p=.263

TERRAIN TWO:

Table of Differences between percent correct identification of Landmarks over sites, including mean percent correct for each Landmark. Column Mean is subtracted from the Row Mean to obtain the Difference Score. [HSD = .1545]

	2	3	4	5	6	7
1: 31.94	-08.34	-22.34*	-06.25	-09.03	02.08	06.94
2: 40.28		-14.58	02.09	00.69	10.42	15.28*
3: 54.86			16.67*	13.89	25.*	29.86*
4: 38.19				-02.78	08.33	13.19
5: 40.97					11.11	15.97*
6: 29.86						04.86
7: 25.						

* p<.05, HSD=15.45

Table of Correlations between Landmarks in Terrain Two (N=24).

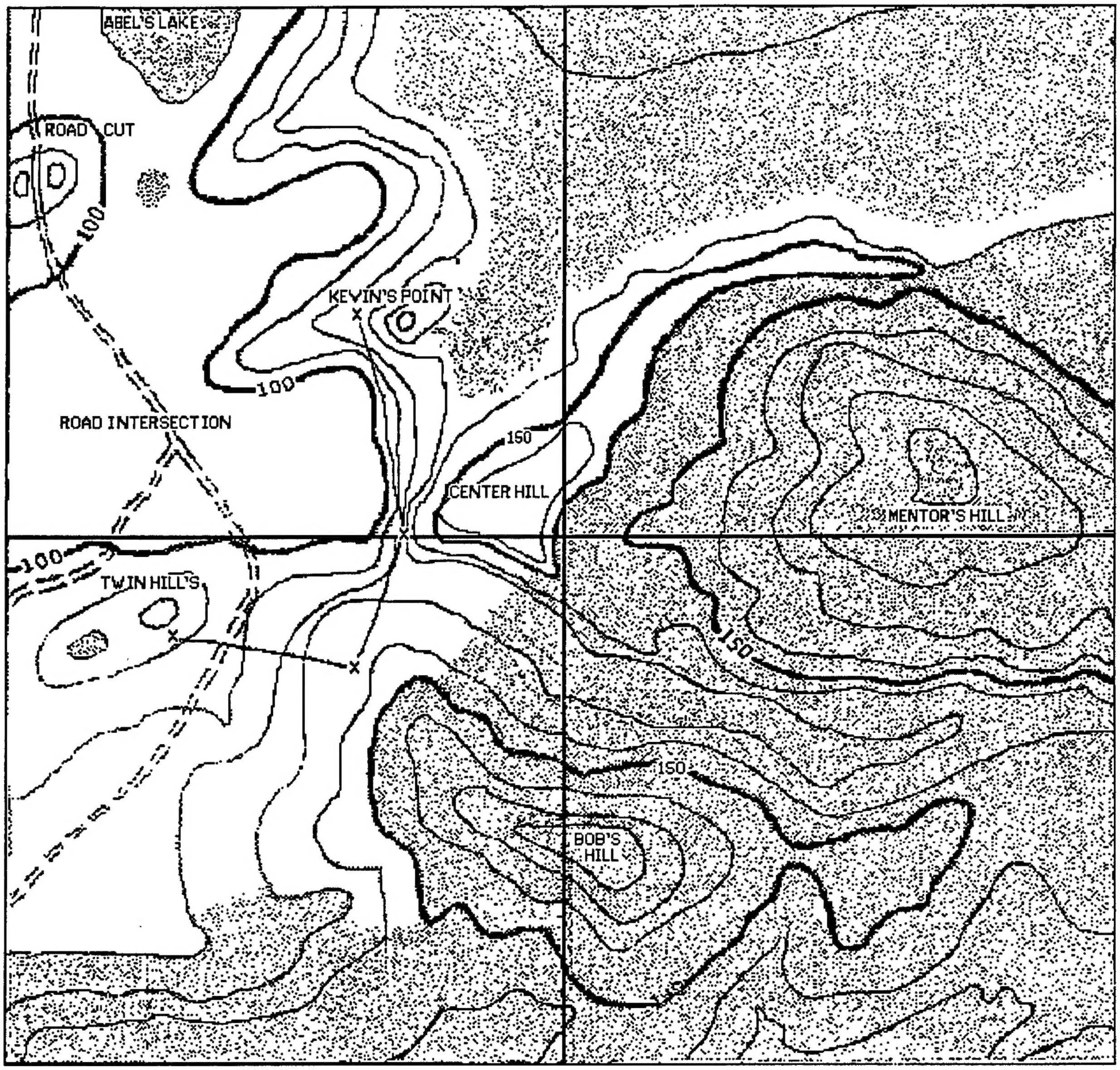
	1: Selby	2	3	4	5	6
2: Torrel	.4129					
	p=.045					
3: Pond	.3704	.5928				
	p=.075	p=.002				
4: McKenna	.3710	.4681	.2192			
	p=.074	p=.021	p=.304			
5: Rd Int.	.0909	.4403	.2291	.3386		
	p=.673	p=.031	p=.282	p=.106		
6: Walter	.1894	.2785	.0707	.4417	.1864	
	p=.375	p=.188	p=.743	p=.031	p=.383	
7: Higley	.4341	.2831	.1158	.4236	.2459	.2154
	p=.034	p=.18	p=.59	p=.039	p=.247	p=.312

Appendix F

Topographical Maps of Experimental Terrains

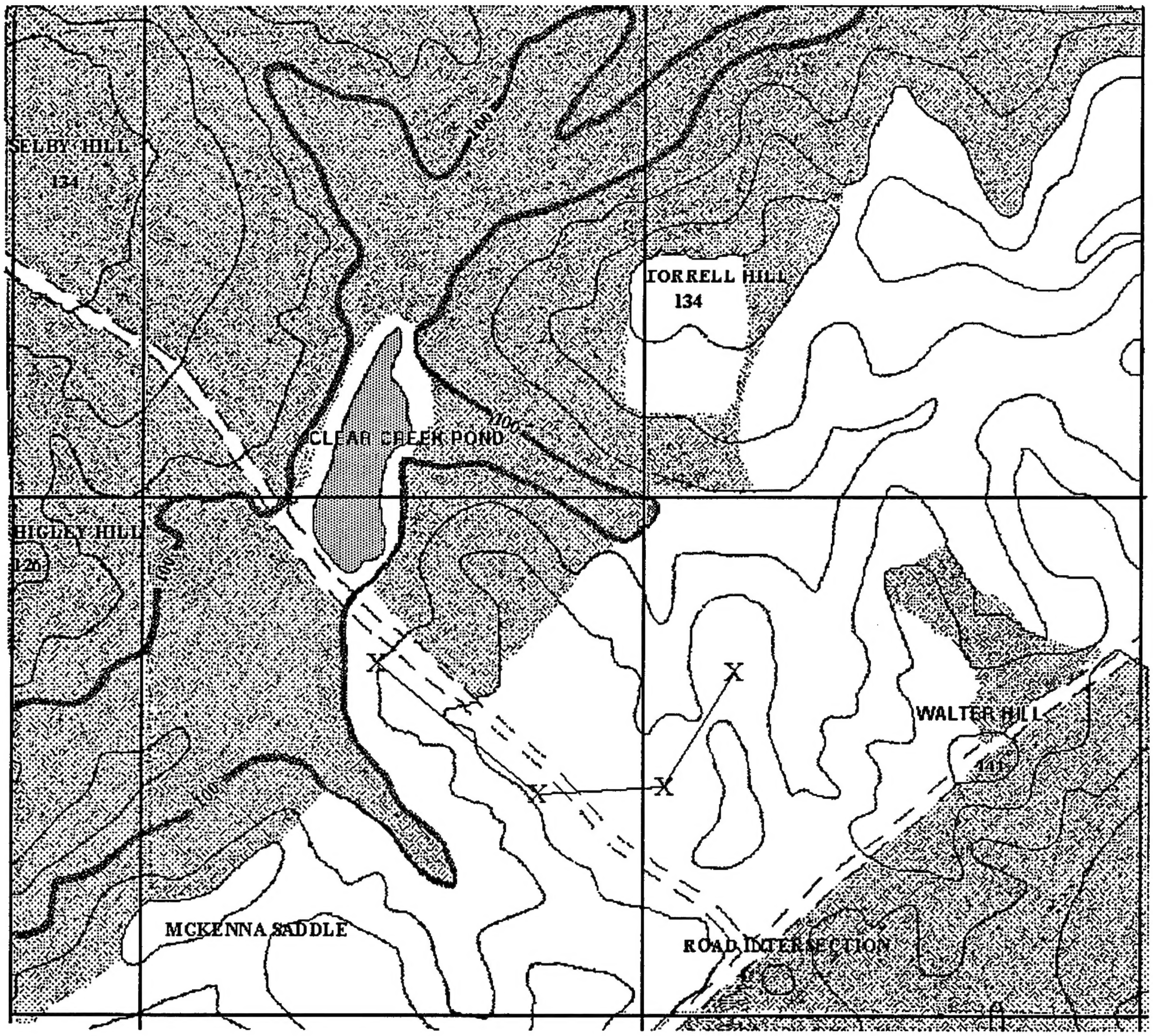
F2

Topographical Map of Terrain One.



The terrain crossing path followed during training is indicated by the connected X's on the map. The starting point was always the X at Twin Hill's. None of the hills in the terrain were over 200 meters in height. The area represented is a two by two kilometer area.

Topographical Map of Terrain Two.



The path followed during training is indicated by the connected X's. The starting point was always at the X south of Clear Creek Pond, and near the road. This terrain matches terrain found at Ft.Benning, near the McKenna MOUT site training area (south and west of McKenna's Saddle). All maps used by the subjects contained a compass (North is up), and a scale (on the maps used, one centimeter equaled one hundred meters, and each contour line is ten meters). The area represented is approximately 2.3 kilometers (east to west) by 2 kilometers (north to south).